Improving Massive Courses with Micro Games

The Effect of Small Serious Games on Student Retention in MOOCs

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Systems Engineering, Policy Analysis and Management**

Faculty of Technology, Policy and Management

by

Jaron Castelijn

Student number: 4173260

To be defended in public on November 1, 2017

Graduation committee

Chairperson	: Prof.dr.ir. A. Verbraeck, Section Policy Analysis
First Supervisor	: Dr. C. van Daalen, Section Policy Analysis
Second Supervisor	: Dr. ir. G. Bekebrede, Section Policy, Organisation, Law & Gaming
External Supervisor	: D. Groen MSc, InThere

Acknowledgements

While designing goal trees, problem statements and causal diagrams for in the serious game, I always found it symbolic that to some extent I was graduating on the same subject as I started my student career with. It turned out that the contents of my first course at TPM were also present in my last work. Not only that, I had to apply this knowledge to do some serious stakeholder management in order to finish my thesis, as I was dependent on many "actors" during my graduation. That being said, stakeholder management is not that difficult if the stakeholders are always willing to help, and for that I am very grateful.

I would first like to thank my graduation committee, starting with Els van Daalen. You did not only help me to find a subject and matching graduation committee, but you always provided the right amount of coaching and constructive feedback to keep me going. I always felt welcome to come by your office, and appreciate the time and effort you put into helping me in one of the many situations I was dependent on other "actors" during my graduation. Next, I would like to thank Daan Groen, for showing me the world of serious game development at InThere, for providing me valuable insights in designing a serious game, and for our friendly and insightful discussions regarding games and many other things. Third, I would like to thank Geertje Bekebrede. It struck me that while we did not meet often, your honest feedback regarding the theoretical framework and the role of the serious game design in my thesis influenced the result greatly. Last, I would like to thank Alexander Verbraeck for not only the feedback itself during the kick-off, mid-term and green-light meetings, but your ability to provide the feedback in a way that made me more enthusiastic about my own thesis.

As I said, I was dependent on many actors during my graduation. I would therefore like to thank my colleagues at InThere for developing the serious game, Fieke Miedema for her help with the CSPD MOOC facilitation, and Jan-Paul van Staalduinen for assisting with the edX integration and data collection.

Furthermore, I would like to thank my parents for their unconditional support, not only during my graduation, but throughout my entire student career. I could have not done this without you.

Lastly, I would like to thank Yvette for helping me, supporting me, listening to me and motivating me during this period, even more than you normally do.

Summary

A new and rising phenomenon in the world of online learning are Massive Open Online Courses (MOOCs). They have been heralded for offering high level education for people with limited high education possibilities, and for promoting lifelong learning, realizing the shift towards a more knowledge based society. MOOCs have also been criticized by many for their low completion rates (Saraguro-Bravo, Jara-Roa, & Agila-Palacios, 2016), which averaged around 6.5% (Jordan, 2014). The low student retention is considered a problem as it requires a significant amount of time, effort, and infrastructure to develop MOOCs, and these are valuable resources for educational institutions (Kim et al., 2016).

In the bulk of literature that analysed why MOOC retention is low and what factors influence it, different approaches are suggested to increase retention, of which quite a few mention serious games as possible method (De Freitas, Morgan, & Gibson, 2015; Gütl, Rizzardini, Chang, & Morales, 2014; Hew, 2016; Meyer, 2014). Suggesting serious games as method to increase low student retention in MOOCs is most likely based on two key-publications of Malone (1981) and Garris, Ahlers and Driskell (2002) who state that serious games are suitable for educational purposes because of their motivational potential.

However, when analysing the results of several meta-analyses on the effectiveness of serious games, there is no clear consensus on this motivational impact of using serious games. (Clark, Tanner-smith, & Killingsworth, 2014; Sitzmann, 2011; Vogel et al., 2006; Wouters, Van Nimwegen, Van Oostendorp, & Van Der Spek, 2013). Sitzmann (2011, p. 513) in particular states that, given that serious games are primarily known for their motivational potential, *"It is ironic that a dearth of research has compared post training motivation for trainees taught with simulation games to a comparison group"*. In this report this knowledge gap is addressed by sharing the results of our experiment on the motivational impact of serious games when utilized to tackle low student retention suffered by MOOCs. The following main research question was drafted:

To what extent can student retention in MOOCs be improved by incorporating a serious game?

Research approach

This research question was answered with a research approach that consists of two steps (1) identifying the theoretical possibilities serious games offer to influence student retention in MOOCs and use these to construct a theoretical framework and (2) test if a serious game that is designed according to this framework has an impact on the student retention when incorporated into a MOOC. The two main research methods used are a literature study for step 1 and an experiment for step 2. The serious game in the experiment was developed by the Dutch company InThere (www.inthere.nl), and designed following their adapted approach from the Triadic Game Design (TGD) method developed by Harteveld (2011). In the experiment the serious game was incorporated into the Creative Problem Solving (CPSD) MOOC, which is hosted on the edX platform. The course was offered by DelftX, an online learning institution affiliated with Delft University of Technology.

Theoretical possibilities for games

The literature study resulted in the identification of four factors that were shown to be related to student retention in MOOCs, and might also be influenced by a serious game. These four factors were *intrinsic motivation*, *extrinsic motivation*, *interactivity* and *perceived MOOC quality.* With these identified factors, a theoretical framework was developed that represented how the serious game could influence student retention in a MOOC, consisting of three parts.

1. The relation between the four factors and student retention, but also the relations between the four factors themselves.
2. The influence of the serious game incorporation on the four factors. The relations between the four factors are not accounted for, as these are already researched in the first part of the theoretical framework.
3. The relation between the serious game and the student retention.

Experiment set up

The experimental design used in this study was a randomized post-test only control group, in which the treatment group, or game group, had access to the serious game, while the control group followed the normal course. The four factors were measured using a construct-item relation, and an Explorative Factor Analysis (EFA) was used to construct the four factor scores. The student retention was represented by *if students completed the course or not*, and *how engaged they were with the course*. Students completed the course if the obtained a grade of at least 60%. Course engagement was measured with the number of assignments and forum posts made per student. The data regarding the student retention stemmed from the edX platform. Several other supportive analyses were conducted, including a small quantitative and qualitative game evaluation.

Results

- Part 1: It was analysed if the four identified factors influenced student retention in this MOOC, by comparing the intrinsic motivation, extrinsic motivation, interactivity and perceived MOOC quality between students who completed the MOOC (N = 39) and students who did not (N = 13). No significant differences were found for any of the four factors. There was a significant moderate correlation found between extrinsic motivation and interactivity.
- Part 2: After analysing if the serious game incorporation had an impact on the four factors, by comparing the four factors scores between the students in the game group (N = 22) and students in the control group (N = 30), it was found that students in the game group had a lower intrinsic motivation.
- Part 3: It was also analysed if the serious game had an impact on student retention by comparing the completion rate, number of assignments made and number of forums posts between the game group (N = 112) and control group (N = 142). No significant differences were found in any of these analyses.
- Supportive: the quantitative game evaluation (N = 20) showed that the students regarded the game as moderately positive regarding all aspects, including the questions related to intrinsic motivation. The qualitative game evaluation showed that the game was both perceived as positive and negative by students, and that the timing of the games seemed responsible for most negative responses.

Conclusions

The results of experiment show that the serious game incorporation did not affect student retention in this MOOC. The lack of impact of the serious game is not surprising considering the fact that the game was specifically designed to influence four factors that were shown to affect student retention in other MOOCs, but did not appear to affect student retention in this MOOC. Furthermore, the intrinsic motivation appeared to be lower in the game group, most likely due to the challenging timeframe in the game. These findings suggest that the extent to which serious games could improve student retention in MOOCs is currently limited, as first the underlying factors affecting student retention in MOOCs must be further researched.

Scientific and societal contributions

This study contributes to MOOC literature regarding student retention by providing an overview of the factors that influence student retention, and by showing that these factors do not yield the same results when analysed in a different MOOC. Furthermore, this study adds to the debate on the motivational capacity of serious games and the debate on how to measure student retention. The study contributes to serious gaming literature by exploring the possibilities serious games offer to motivate students in online learning, and, as being one of the first studies to do this, by showing how a serious game could be integrated into a MOOC. The findings of this study can be used to improve methods aiming to increasing student retention in MOOCs, therewith contributing knowledge to the improvement of free-access online learning.

Limitations

First, there were a low number of responses to the questionnaires, decreasing the validity of the statistical analysis. The validity of the factor analysis was decreased due to the limited number of questions that could be incorporated in the questionnaire. This led to the limitation of measuring *interactivity* with only student-teacher interaction. Furthermore, it was not possible to measure the influence of the serious game on the student engagement in the course, meaning only conclusions could be made on the serious game influence on the completion rate. As the time and budget available for the serious game development was limited, some factors were only accounted with textual encouragements instead of game mechanics. A self-selection bias of mostly students who completed the MOOC filling in the questionnaires, meant that the results of the serious game influence on the four factors could only be generalized to students who complete the MOOC. Lastly, as only one MOOC with a tailored made serious game was researched, the external validity to other MOOCs and serious games is low.

Recommendations for future research

There were several recommendations made for future research: (1) Re-run this experiment with an improved questionnaire, a larger sample (while actively preventing a low response from drop-outs), including student-student interaction, and including the effect the serious game on student engagement; (2) validate the exploratory factor analyses used in this study with a Confirmatory Factor Analyses in another similar MOOC; (3) use a similar serious game to this study in another MOOC; and (4) analyse the same MOOC with an improved serious game design, possibly a serious game focused on increasing interaction between students, persuading students to start the MOOC, or being appealing to users with diverse backgrounds.

Table of Contents

List of tables

List of figures

1. Introduction

Ever since the introduction of the first Massive Online Open Course (MOOC) in 2008, the popularity of MOOCs has been increasing. MOOCs offer free access to university level information and education for a large number of people via an online learning environment. They have been heralded for offering high level education to people who previously had no access to it and for promoting lifelong learning, realizing the shift towards a more knowledge based society. Currently there are close to 60 million students enrolled in almost 7000 courses offered by more than 700 universities (Dharwal Shah, 2016). Most MOOCs use quizzes, online assignments and examinations to provide interactive educational content (Tan, 2013), and interactions with other students are accounted for via discussion fora (Breslow et al., 2013). Due to the steep increase of enrolled students in the past five years, currently a highly competitive environment exists where courses, platforms, and e-learning strategies are constantly adapted to outcompete each other (Freire, Del Blanco, & Fernańdez-Manjón, 2014).

However, MOOCs have also been criticized by many for their low completion numbers (Saraguro-Bravo et al., 2016). The results found by Jordan (2014) showed that the average completion rate of a MOOC is only 6,5%. This high drop-out of students during the course is a well-known characteristic of MOOCs and has even been named by Clow (2013, p. 186) as the “The Funnel of Participation”.

The low student completion is considered a problem as it requires a significant amount of time, effort, and infrastructure to develop MOOCs, and these are valuable resources for educational institutions (Kim et al., 2016). The resources are invested for a diverse set of reasons: strategic growth of university, marketing possibilities, strategic collaboration with other educational institutions, organic growth and evolution, response to learner requests, to get valuable insights via Learner Analytics and for educational enhancement (White, Davis, Dickens, León, & Sánchez-Vera, 2015). In addition, some MOOCs provide the possibility to pay for the MOOC to acquire an official certificate of completion, which means that low student retention could lead to missed potential income. Lastly, on a nobler note, the low student retention decreases the effectiveness of MOOCs to provide world-wide free access to university level knowledge.

1.1 Problem exploration

The low student retention in MOOCs spawned several studies investigating the subject. Some studies investigated the low retention rates by looking at the intentions students had for enrolling and found that a large group of enrollers did not have the intention to complete the MOOC in the first place (Belanger & Thornton, 2007; Gütl et al., 2014; Kizilcec, Piech, & Schneider, 2013; Wilkowski, Deutsch, & Russell, 2014). Other studies focused on why students enrolled, and found that students had both intrinsic and extrinsic enrolment motivations (Belanger & Thornton, 2007; Davis, Dickens, Leon, Del Mar Sanchéz Vera, & White, 2014; Hew & Cheung, 2014; Xiong et al., 2015). When looking at course level variables Jordan (2014) found a negative correlation between student retention and course length, but this finding was not supported by for example Adamopoulos (2013). Findings regarding student level variables like age, gender, prior experiences also differed per study (Adamopoulos, 2013; Greene, Oswald, & Pomerantz, 2015; Wilkowski et al., 2014). In the bulk of literature that analysed why MOOC

retention is low and what factors influence it, different approaches are suggested to increase retention, of which quite a few mention serious games as possible method (De Freitas et al., 2015; Gütl et al., 2014; Hew, 2016; Meyer, 2014) for similar or different reasons. However, only two studies were found that proceeded to incorporate a serious game into a MOOC, and both yielded promising results regarding student retention (Bernaert, Crepon, & Dhorne, 2015; Romero & Usart, 2013).

Introducing serious games

Using serious games as method to increase low student retention in MOOCs probably stems from two key-publications of Malone (1981) and Garris, Ahlers and Driskell (2002) who state that serious games are suitable for educational purposes because of their motivational potential (Sitzmann, 2011). As the students apparently lack motivation to finish the MOOCs, it makes serious games interesting to address the student retention problem. Serious games have been used throughout history for different purposes, but most research regarding serious games have focused on serious games with the primary purpose to educate, using serious games as a teaching method.

The results of several meta-analyses on the impact of simulation, serious, or digital games show that game-based-instructional methods are indeed preferred over traditional teaching methods regarding learning outcomes (Clark et al., 2014; Sitzmann, 2011; Vogel et al., 2006; Wouters et al., 2013). However, there is no clear consensus on the motivational impact. Wouters et al. (2013) give three possible explanations for this, one of which is the favouring of the instructional design over the entertainment design, resulting in an instructive but less engaging game. This is interesting for this research, as not necessarily an educational, but a "retention-increasing" serious game is needed.

An important note is that these meta-analysis all compare the use of game-based teaching methods with traditional teaching methods in a face-to-face learning environment. The two studies that research the serious game incorporation into a MOOC both did this differently with varying quality. Romero and Usart (2013) conducted a qualitative research on the student and teacher experiences when serious games were used as main part of the course. They found that the students highly appreciated the serious games, but did not look at a possible effect on student retention. Bernaert et al. (2015) did quantitative research into the effects of a serious game as main instruction method in a MOOC and as evaluation method at the end of each week. They reported a high completion rate for their MOOC (31%) and stated that students found the game helpful, a big part even found it the most enjoyable part of the course. However, as none of the researchers using an experimental design, it is unclear to what extent the high completion rate can be attributed to the presence of the serious game.

Problem statement

MOOCs experience low student retention, leading to valuable resources being ineffectively spent to achieve university goals, less income being generated and university level education being less effectively distributed around the globe. It is unclear what possibilities serious games provide to solve this problem, as not much research has been conducted so far.

1.2 Research goal

From the problem exploration it can be derived that serious games have the potential to tackle the low retention problem suffered by MOOCs, but so far serious games were mostly designed for educational purposes in traditional face-to-face learning environments. Additionally, one may wonder what the effect would be if not the educational but the motivational design would be favoured, as indicated by Wouters et. al. (2013). Taking these considerations into account, the following research goal was formulated:

Research goal

Getting insight into the possibilities serious games offer to increase student retention in MOOCs.

1.3 Relevance of study

Scientific relevance. The scientific relevance of this project is mostly related to the identified knowledge gap, by contributing knowledge on the capability of serious games to improve student retention in MOOCs. Note that this contributes to both the MOOC field of literature and serious game field of literature. Furthermore, as this research builds further upon earlier findings regarding low student retention in MOOCs, this study contributes scientific knowledge by somewhat replicating previous findings in different contexts.

Societal relevance. The societal relevance of this research is contributing knowledge to solve the low retention problem experienced worldwide. Experimenting with one of the suggested methods in recent literature that could improve student retention in MOOCs, gives valuable insight into the possibility serious games offer MOOC facilitators to improve their courses, as well as new insights that could inspire different approaches or improve other methods.

1.4 Research questions

As shown in the previous sections, this research is concerned with assessing the possibilities serious games offer to increase student retention in MOOCs. Although there is a large exploratory component due to the limited amount of research on the retention improving capabilities of serious games in MOOC environments, the scientific and societal communities would benefit most from testing the possibilities serious games offer to increase student retention in practice, which is why the following main research question of explanatory nature was formulated.

Main research question

To what extent can student retention in MOOCs be improved by incorporating a serious game?

1.4.1 Sub-research questions

To be able to answer "to what extent" serious games can improve student retention in MOOCs, it is first important to understand what influences student retention in MOOCs and how this could be influenced by a serious game. This enables to grasp the possible ways a serious game could influence student retention more clearly, and therewith serves as theoretical framework for the design and evaluation of the serious game. During the initial

literature study in Section 1.1, it became clear that the field of serious games and MOOCs are both new and grow fast. Nevertheless, not much research was found that combines the three key concepts used in this research: serious games, MOOCs and student retention. This showed that no theoretical framework concerning the use of serious games for improving student retention in serious games exist and had to be developed first. The development of this theoretical framework is accounted for by answering sub-research questions 1 and 2.

However, with the development of the theoretical framework the main research question is not answered yet, as the explanatory nature is not addressed. For this, the effects of a serious game incorporation into a MOOC had to be evaluated. First the theoretical framework is evaluated by answering research question 3 and 4. Finally, the effect of the serious game incorporation into a MOOC on student retention is addressed by answering research question 5.

Answering research question 5 shows *if* a serious game leads to improved student retention, and the answers of research question 1 to 4 provide the theoretical framework to get more insight into *how* this result was or was not realized. When combined, the answers can be used to evaluate *to what extent* serious games can be used to increase student retention in MOOCs, therewith answering the main research question.

Theoretical framework development research questions

1. What factors could influence student retention in MOOCs?
2. Which factors influencing student retention in MOOCs could be affected by a serious game?

Theoretical framework evaluation research questions

3. To what extent do extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality influence student retention in MOOCs?
4. To what extent does a serious game affect extrinsic motivation, intrinsic motivation, interactivity, and perceived MOOC quality?
5. Is there a difference in student retention when a serious game is incorporated into a MOOC?

1.5 Research approach

Looking at the research questions, this study was divided into two parts: the theoretical framework development (research question 1 and 2) and theoretical framework evaluation (research question 3, 4 and 5). The theoretical framework development is, as the name suggest, of theoretical nature and is researched with a literature study. The theoretical framework evaluation is of practical nature and is researched with an experiment. As these two methods constitute the main two research methods, they are more extensively elaborated upon in Sub-section 1.5.1 (literature study), and Sub-section 1.5.2 (experiment).

1.5.1 Literature study

The two main strands of the literature study were MOOC-related literature and the serious games related literature, with literature concerning student retention being represented in both, and sometimes with additional searches.

MOOC literature study

Literature concerning MOOCs was searched using the databases of Scopus, IEEExplorer, JSTOR and Google Scholar. The initial search terms for this search were "Massive Online Open Courses" AND "MOOCs". To narrow down this initial search, the focus was on peer-reviewed publications written in English, such as journal articles, conference proceedings and books. This provided a range of articles about the motivation of participants, the motivation of facilitators, the predictability of drop-outs and different methods to decrease the drop-outs. At first an article that could provide an overview of the existing MOOC literature was looked for by adding "Literature review" AND "Systematic study". This led to the finding of an article that provided an overview of all MOOC related articles from the first MOOC in 2008 till 2012 (Liyanagunawardena, Adams, & Williams, 2013). The article showed a steep increase in MOOC related articles. In 2008 and 2009 only one scientific article was published, but this grew rapidly to seven in 2010, ten in 2011 and twenty-six in 2012 (Liyanagunawardena et al., 2013). By conducting a search in Scopus with the terms "MOOCs" and "Massive Open Online Courses" the graph in Figure 1 was obtained. The graph resembles the trend mentioned by Liyanagunawardena (2013) of papers published prior to 2013 and shows the continuation of that trend in the last years.

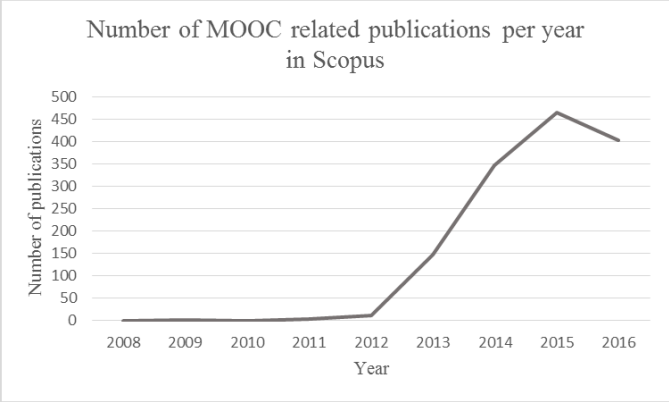


Figure 1: MOOC related article publications per year in Scopus

The article of Liyanagunawardena (2013) provided a good overview of the published literature prior to 2013, but for the time period after 2012 an additional systemic literature review was required. The results were found to be too broadly oriented, due to the recent uptake of literature, and the topics were varied and unclearly defined. Also, it was found that due to different definitions of key terms, authors often did not reference to similar publications in the years before. Therefore, the literature review by Veletsianos and Sheperdson (2016) was used to get an insight in the literature from 2013 till 2015. They found 5 research strands: student-focused (popular topics were completion and retention and learner subpopulations), design focused (design, creation and implementation of MOOC themselves), context and impact of MOOCs (usefulness, perceptions on MOOCs and economic impacts), instructor-focused (their experiences with MOOCS) and other. The findings are presented in Table 1, take note that the total frequency exceeds 100% as research strands were often combined. This study is mainly concerned

with student-focused research, which is why for the main body of MOOC literature the terms “Retention” OR “Completion” OR “Drop-out” OR “Attrition” OR “Persistence” were added. At some points findings in other strands were used as well. In those cases, the MOOC research terms were expanded with relevant terms.

Table 1: Research strands in MOOC literature 2013-2015 (Veletsianons & Shepherdson, 2016)

Research strand	Frequency (%) of total papers
Student-focused	83.6
Design-focused	46.4
Context and impact	10.9
Instructor focused	8.2
Other	9.8

Serious gaming literature study

Literature concerning serious games was searched using the databases of Scopus, IEEExplorer, JSTOR and Google Scholar. As using just “serious games” as initial search terms would yield too many results, it was complemented with “definition”. This led to the finding of *Serious games – an overview* (Susi, Johannesson, & Backlund, 2007) amongst other literature, that pointed out the discussions and debates concerning serious games. One of these discussions concerned the definition of a serious game and its relation to other concepts. The serious gaming definition literature was reviewed by means of backward references searching.

Another big part of the serious gaming literature concerned the educational effects of serious games as opposed to traditional teaching methods. At first specific effects of serious games on student retention was searched for by adding the inclusion criteria (“Retention” OR “Completion” OR “Drop-out” OR “Attrition” OR “Persistence”) to (“Serious Games” OR “Simulation games”). The findings covered many types of research concerning serious game design, effects and assessment studies. Instead of an extensive literature review on the effects of serious games it was decided to review several meta-analytic examinations to assess the findings in the past decade on the effects of serious games on education and motivational aspects. To find other meta-analyses than the first one found by Sitzmann (2011), forward referencing was applied.

1.5.2 Experiment

As this research aims to test the effect of a serious game incorporation in a MOOC, an experiment was conducted. For the experiment being able to imply a causation between the serious game incorporation and the effect on student retention, it is required to make two randomized groups with different treatments, and then analyse the results with several statistical methods.

The student retention problem is also experienced by the Creative Problem Solving and Decision Making MOOC offered by the DelftX. Depending on when you count students as “started”, the completion rate for this course last year varied from 0,85% to 8%. Naturally the facilitators of the course want to increase this number as it can be seen as an indicator of how successful the course is, but also to mitigate the disadvantages of low student retention mentioned in the problem statement. For this MOOC it was possible to get access to MOOC and student data via

the TU Delft Extension School. The Extension School is the supporting organisation that helps, advises and guides lectures in online education (Extension school, 2017). Important to note about this experiment is that the CSPD MOOC already started before this research was conducted. However, with the possibility to intervene with the MOOC and the access to the MOOC data, this MOOC is deemed appropriate for the experiment that was needed to test the theoretical framework.

Note on game development method

As the experiment requires a serious game to be developed this section briefly elaborates on the design approach. There is large body of research available on serious game design, but as the main focus of the research is not on serious game design, a practical game development approached was favoured over a theoretically founded one. Especially since the short time span in which the game had to be designed, developed and integrated into an online platform, required a short development time. As the researchers themselves did not possess the ability to develop a digital game, it was therefore chosen to have the serious game externally developed by InThere, a Dutch company that is specialized in the short-time developing of small serious games (microgames) to accelerate change and training project in companies. Microgames are small serious games with a single learning goal, which are played in sessions of about 15 minutes.

The InThere game development approach is an adaptation of the Triadic Game Design (TGD) method developed by Harteveld (2011), of which the process development time is greatly reduced based on experience from the private-game development sector. Furthermore, with InThere's "Whitebox", a game hosting platform that can be used to collect game data, the process of collecting game data could also be simplified. InThere's game development process consists of roughly three phases, the Gamestorm, the game design, and the game development. The Gamestorm is the most important aspect of the development approach and is therefore briefly elaborated upon.

The Gamestorm is a method developed by InThere in which the clients are invited to a workshop to intensely discuss the experienced problem. By providing a structured method during this workshop with all actors present, the ambiguity of a problem that is normally experienced when communicating with several people from the client's side through several forms of communication, is reduced. Furthermore, an agreement on the problem description at the end of the Gamestorm serves as a "mental" contract to prevent clients form changing their wishes halfway through the design, which delays the development process. Lastly, the Gamestorm serves as method to make the problem owner also feel responsible for the game design, ensuring their cooperation and enthusiasm. The focus of the Gamestorm is mostly with the *Reality* and *Meaning* aspect of Triadic Game Design approach developed by Harteveld (2011), but rough ideas for the *Play* (game) aspect are often touched upon. After the Gamestorm, the results regarding the Reality, Meaning and Play were used as input for a preliminary design. This preliminary game design is normally presented to the clients after two weeks. When the clients agree to the preliminary design, the development phase is initiated.

1.6 Thesis outline

This chapter is concluded with an outline in Figure 2, which provides an overview of the research questions, research methods and products per chapter. The research questions are formulated on the left side of the figure. The research methods are depicted with the white rectangles, the products with blue ovals. The arrows indicate the relation between the research methods and products.

In the Figure it can be seen that the products of Chapter 1 are the problem statement, research goal and research questions. The research questions guided the literature study in Chapter 2, to enable the development the theoretical framework. The theoretical framework is represented by factors that, according to the literature study, influenced student retention in other MOOCs and might be influenced by a serious game. The theoretical framework is operationalized into workable hypotheses that could be tested in an experiment described in Chapter 3. As the experiment required a serious game to be incorporated into the MOOC, a serious game was developed using the Triadic Game Design adaptation by InThere, taking into account the identified factors from the theoretical framework. In Chapter 4, several statistical analysis methods were used to analyse the data from the experiment. The experiment provided quantitative data that could be used to test the hypotheses, and a mixture of quantitative and qualitative data that could be used to support these findings and provide more context for interpretation. In Chapter 4 the data from the experiment is analysed using several statistical methods. The results are discussed in Chapter 5, before the conclusions are drawn in Chapter 6.

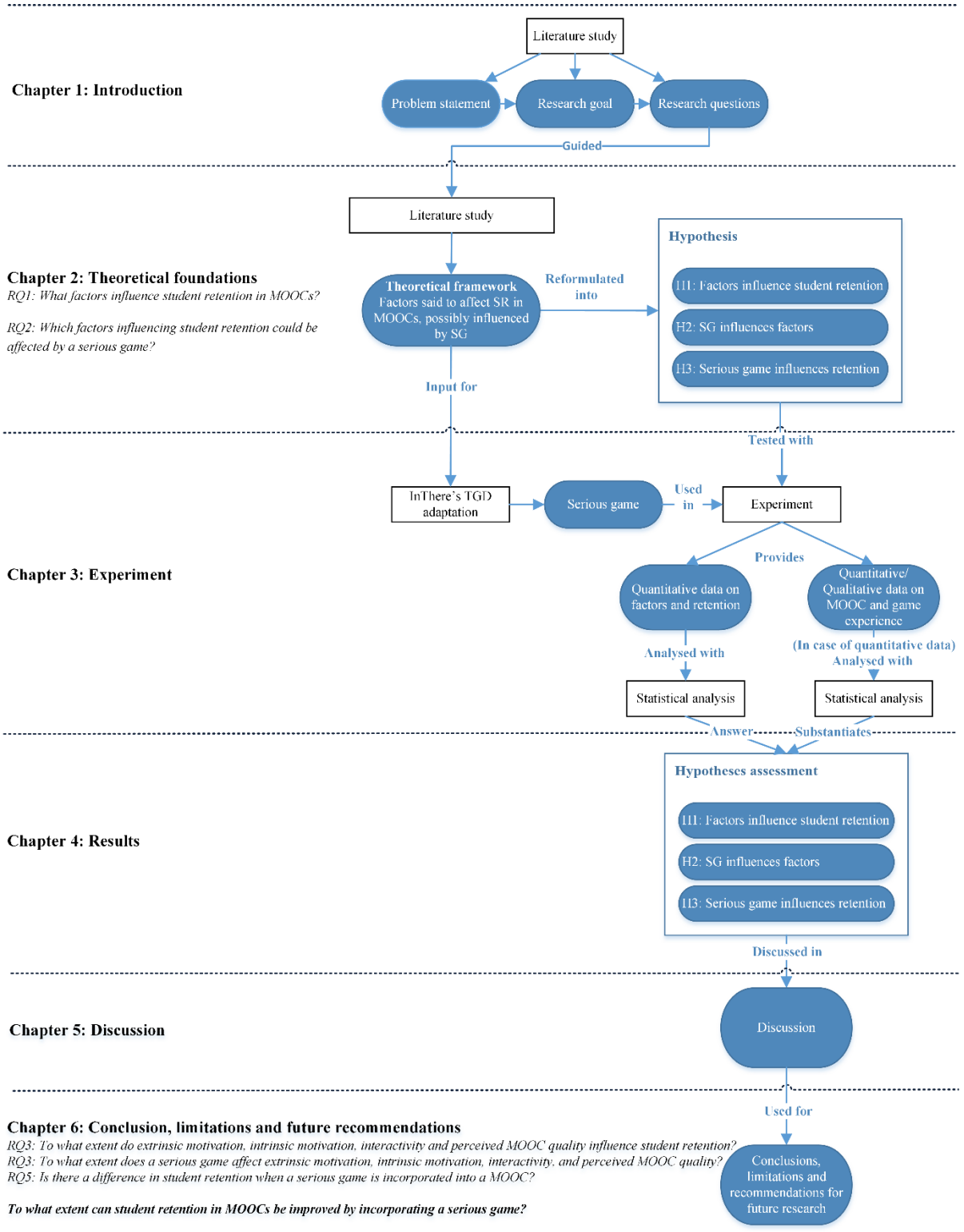


Figure 2: Outline of thesis with research methods, products and relations

2. Theoretical foundations

The aim of this chapter is to develop a theoretical framework represented with hypotheses that can be tested with an experiment, by answering the first two research questions:

1. *What factors influence student retention in MOOCs?*
2. *Which factors influencing student retention could be affected by a serious game?*

First, the three key concepts in this study, serious games, MOOCs and student retention, are defined in Section 2.1. These definitions demarcate the related literature study in Section 2.2, of which the findings are used to answer research question 1 (Sub-section 2.2.1) and research question 2 (Sub-sections 2.2.2 and 2.2.3). Lastly, the answers to the research questions are used to develop a theoretical framework represented with hypotheses in Section 2.3. To keep provide a clear overview of the theoretical foundations, the chapter structure is illustrated in Figure 3.

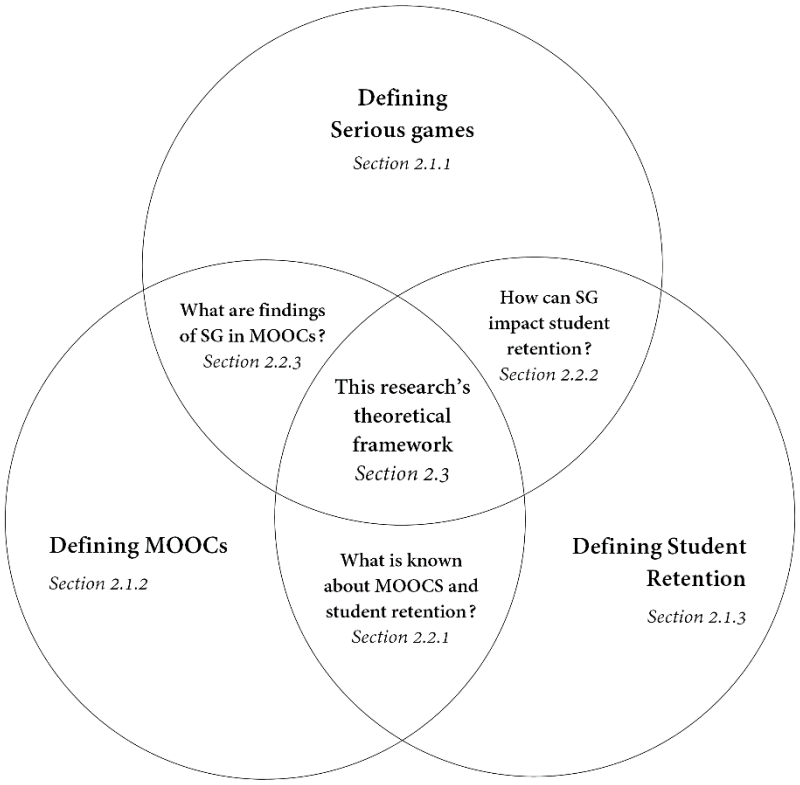


Figure 3: Theoretical foundations chapter structure

2.1 Positioning this research

In the available literature surrounding serious games, MOOCs, and student retention, a variety of names and definitions are used to describe these concepts. Untangling this web of research was not only required to position this research, but to interpret possible relevant findings from other studies as well. First the definition of serious games is described in Sub-section 2.1.1, followed by MOOCs in Sub-section 2.1.2 and student retention in Sub-section 2.1.3.

2.1.1 Defining serious games

When reading game-related literature, you come across a variety of names to describe the same or different concepts: virtual experiences, simulations, social impact games, practiceware, game-based learning, immersive learning environments, educational simulations, serious games, sims, epistemic games, edutainment, edugaming, simulation/gaming, simulators, persuasive games, virtual training environments, and so on (Harteveld, 2011). Harteveld states that the names *simulations* and *serious games* are the two most used names in game-based literature.

One of the first authors to name and define serious games is Abt (1977, p. 9). He defines serious games as having "an explicit and carefully thought-out educational purpose and are not intended to be played primarily for entertainment". In this definition, the educational purpose is specifically mentioned as prerequisite to be a serious game. However, the purpose of a serious game does not necessarily have to be of an educational nature to be considered serious. For example, the game *Foldit* was used by researchers to have people create accurate protein structure models, instead of having this done by computational models (Cooper S et al., 2010). Since the main purpose of this game is neither educational nor entertaining, it cannot be considered a serious game according to Abt's definition, while it clearly has a serious purpose. Michael and Chen (2005, p. 21) altered the definition of Abt to what they call "the simplest definition of serious games: games that do not have entertainment, fun or enjoyment as their primary purpose". Michael and Chen do add that entertainment games that are used for other purposes, can also be considered serious games. This definition differentiates serious games from entertainment games, without detailing a specific purpose, making it applicable to many forms and types of games.

The term serious games is also criticized. Harteveld dislikes the term serious games as it implies that entertainment games cannot be serious business as well. Furthermore, Harteveld (2011) and Susi, Johannesson and Backlund (2007) note that the term "serious game" can be confusing as the two constituents seem mutually exclusive. How can something be serious and a game at the same time? Whether or not this is possible depends on how you define a game, which sparks another still going debate. The first definition of a game by the Oxford dictionary is "an activity that one engages in for amusement or fun", but the second definition is "a form of competitive activity or sport played according to rules." While the first definition is clearly not compatible with Michael and Chen's definition of serious games, the second one is. In the book of Salen and Zimmerman (2004, p. 96), after they compare eight much cited definitions of games, the authors define a game as: "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome." The term *conflict* is used to describe a contest of power with a game system or with other players. Although Salen and Zimmerman specifically state that a conflict could be violent or non-violent, the term is easily interpreted as violent and the term *contest* is

more preferred in this study. Looking at this definition, the second definition of game in the Oxford dictionary seems to resemble the definitions of authors in game-based literature more closely, weakening the criticism that the constituents in the term serious game are mutually exclusive as it appears that in scientific literature a game is not necessarily considered an activity that one engages in for amusement or fun.

The alternative to the term serious games is *simulations* or *simulation games.* Simulation games can be defined as: "A special type of model that uses gaming techniques to model and simulate a system. A gaming simulation is an operating model of a real life system in which actors in roles partially recreate the behaviour of the system" (Duke & Geurts, 2004). Harteveld (2011) mentions that he dislikes the term "simulations", because it relates too much to dynamic graphical calculators that only require human input at the beginning of a simulation. He also mentions that the term simulation games is already used to describe an existing genre of games, of which the *Microsoft Flight Simulator* series is a famous example. Another reason not to use the term simulation or simulation games is that it suggests that the game represents the real world to a high degree, while this is not always necessary in order to achieve the purpose of a game. If the purpose of the game is to increase hand-eye coordination, very abstract games could suffice. This is also the case for the serious game that is designed in this study. This game has the purpose to increase student retention, for which a high degree representation of the real world is not strictly necessary. Therefore, simulation games are seen as a sub-set of serious games, which are particularly useful for educational purposes concerning complex real-life situations. Since the primary purpose of the games in this research is not to educate but to motivate, the more high-level term "serious game" is preferred. It defined by altering and combining the definitions of Salen and Zimmerman (2004, p. 96) and Michael and Chen (2005, p. 21):

> *"A system in which players engage in an artificial contest, defined by rules, that results in a quantifiable outcome, and that does not have entertainment, fun or enjoyment as its primary purpose".*

Note on Gamification

A distinction has to be made in this study between gamification and serious games, since both terms are sometimes used interchangeably throughout available literature. Without a clear distinction it was difficult to demarcate the literature study on serious games in Sub-section 2.2.2 and 2.2.3. There exist several quite similar definitions of gamification, of which "the use of game elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011, p. 2) is broadly cited. Within this definition, game-elements are not further defined than "elements that are characteristic for games", which leaves a lot of room for interpretation.

Currently, gamification is mostly associated with adding points, levels, trophies, badges, achievements, virtual goods, and leader boards to content (Simões, Redondo, & Vilas, 2013). However, Kapp (2012) emphasises that the real power of gamification lies with other game elements: engagement, storytelling, visualization of characters, and problem solving. He defines gamification as follows: "Gamification is using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems." (Rice, 2012, p. 10).

Interestingly, the distinction between serious games and gamification cannot be clearly made on the basis of these definitions. Martí-Parreño et al. (2016) state that the key difference between gamification and game-based learning methods, of which they consider the use of serious games to be one, is that for gamification a (video)game is not required. But saying that there are "not required" actually implies that using a serious game can still be considered gamification. Kapp (2012) states this more explicitly, by saying that serious games are a specific sub-set of the meta-concept of gamification, since building a game based on the content to be learned is actually the gamification of the content. However, using the definition of a serious game as presented above, in this research the difference between gamification and serious games is explicitly made on the fact that serious games require an *artificial* contest, which does not comply with adding game elements to *real* non-game contexts.

2.1.2 Defining MOOCs

Although the topics, participant numbers and other characteristics may differ across MOOCs, the name Massive Open Online Courses is not as debated as "serious games" and is used throughout the literature. MOOCs are defined by Jansen and Schuwer (2015, p. 11) as follows:

> *"MOOCs are online courses designed for large numbers of participants, that can be accessed by anyone anywhere as long as they have an internet connection, are open to everyone without entry qualifications, and offer a full/complete course experience online for free"*

MOOCs are often seen as a new form of online learning or e-learning (Leire, McCormick, Richter, Arnfalk, & Rodhe, 2016). The definition of online learning is debated, as well as its connection to or difference with e-learning and distance education (Moore, Dickson-Deane, & Galyen, 2011). This makes it difficult to assess the applicability of findings in studies regarding serious games prior to the introduction of MOOCs, to this study. As this constitutes the main part of serious gaming literature, this has to be taken into account when for example reviewing the meta-analyses on the effects of serious games on motivation and learning.

2.1.3 Defining student retention

Student retention is a much studied topic in educational studies and spans publications over seventy years (Berge & Huang, 2004). However, as with serious games and online learning, there exists no one definition of student retention. On top of that, authors often do not explicitly define what is meant with the term retention, and sometimes switch between several terms within one publication. For example, De Freitas et al. (2015), initially seem to use *retention* as an overarching term to describe the ability of the course to retain students and measures it with completion rates, but later she states that "*unfortunately, the potential for its use for improving retention and completion rates has yet to be quantitatively or qualitatively tested" (page 468)*, suggesting it are two different things.

Xiong et al. (2015) measure retention as the number of days between the start of the MOOC and the last day of activity by the student. They appear to focus more on student activity than course completion, as other authors often seem to do. For example, Perna et al. (2014), measure the retention *rate* when discussing retention, defined as "the number of those who accessed a lecture in the last module divided by the number who accessed a lecture

in the first module". Not all authors look solely at completion rates, Hone and El Said (2016) measured "learner retention" by asking if they completed the MOOC and if not when they dropped out, how many exercises the student completed and how much content they watched or read. They seem to argue that student retention is about both student engagement throughout the course and course completion, but do not define this explicitly.

The way Hone and El Said measure retention is also in accordance with the increasing number of authors who argue that completion rates are not a relevant metric to measure course success, as they do not take into account the intentions and enrolment goals of the students (Greene et al., 2015; Koller, Ng, & Chen, 2013; Wilkowski et al., 2014). They argue that a large part of students do not intend to finish the course from the beginning, which should be taken into account when measuring course completion. For example, Green (2015) differentiates between completion, partial completion and drop-out, to mitigate the effect of people that drop-out belonging in the same category as people that stop in the last week. Taking these considerations into account, the definition of Berge and Huang (2004, p. 3) is chosen to define retention in this study as it does not solely focus on completions rates, but also on engagement throughout the course.

"Retention is the continued student participation in a learning event to completion, which in higher education could be a course, program, institution, or system"

2.2 Related literature

In this section the related literature concerning serious games, MOOCs and student retention is discussed. First, literature regarding MOOCs and student retention is analysed in Sub-section 2.2.1. Second, the literature regarding serious games and student retention is discussed in Sub-section 2.2.2. Third, available literature that combines serious games and MOOCs is discussed in Sub-section 2.2.3.

2.2.1 What is known about MOOCs and student retention?

The published literature concerning MOOCs and the low student retention has spawned much research to explain this phenomenon. Some authors categorize them by the focus on the phase of the MOOC (registration, activity, completion) (Hew, 2016) or distinguish between student and course level factors that affect student retention (Xiong et al., 2015). This research combines both ways of structuring. First, literature concerning the registration phase is discussed. The registration phase is concerned with the students' intentions and motivation when enrolling. Then, the activity phase is analysed, in which student-level and course-level predictors for course completion are discussed and findings from engagement and completion in online learning are compared to more recent findings in MOOC learning.

Registration phase

From the first journal publication about MOOCs till the end of 2012, no research was conducted on why students enrolled in MOOCs (Liyanagunawardena et al., 2013). The following year, Kizilcec, Piech, and Schneider (2013) looked at engagement patterns and reasons for enrolling to identify the different types of students that enrolled. They defined four categories: completers (students who completed most assessments), auditing (mostly watched

video lectures and did few assessments), disengaging (engaged with first assessments, but then decreased), and sampling (briefly explore content or watch one video). This classification complies with the one of Wilkowski, Deutsch, & Russell (2014) (no-shows, observers, casual learners, and completers), who categorized the students based on their stated intention. Interestingly, they found that only 52.2% of the registrants intended to complete the goal, while 47.8% did not intend to finish the course in the first place. Belanger & Thornton (2007) also found differences amongst student when enrolling for a course. They found that 72% of the students reported they intended to get a certificate, while 28% did not enrol to complete the course. Gütl (2014) also reports that only 23% of the registrants intended to complete the course from the outset. This can explain a part of the low completion rates that MOOCs suffer. The findings regarding the student types and intentions are grouped and presented in Table 2, with an "x" indicating if an author mentioned the student type.

Table 2: Structured findings regarding student types and enrolment intentions

	Kizilcec et al., 2013	**Wilkowski & Russell, 2014**	**Belanger & Thornton, 2013**	**Gütl et al., 2014**
Student types				
Completers	x	x		
Sampling, casual learners	x	x		
Auditing, observers	x	x		
Disengaging	x			
No-shows		x		
Intentions				
Intention to complete		52%	72%	23%

To explain the low completion rates for students who did intend to complete the course, a more detailed look was given at the motivations for enrolling. The enrolment motivations are categorized in *extrinsic*, *intrinsic* and *other*. According to Xiong et al. (2015), extrinsic and intrinsic motivations are clear predictors of student engagement and retention in MOOCs. Intrinsic motivation entails pursuing a task for the satisfaction, engagement or interest the task itself might provide, while extrinsic motivation entails pursuing a task for purposes beyond the task, for example, for payment or to earn a credential (Xiong et al., 2015).

Belanger and Thornton (2007) identified two reasons to enrol that are mostly intrinsic: for fun or entertainment. They also identified which particular motivations participants had to complete the MOOC, which were mostly extrinsic: formal recognition of accomplishment, professional development, and as a supplement to a credit bearing course. Other reasons to enrol were: interaction with students, to support lifelong learning or to understand the subject, and social experience and intellectual stimulation, convenience for people that have limited options to access education, or to experience and explore online education.

White (2015) found that students sign up for an intrinsic reason; because MOOCs satisfy interest, for several extrinsic reasons; MOOCs are useful, to update knowledge and improve resumes, enable learning with the best, and provide professional development and lifelong learning, and other reasons; because MOOCs are free and open and they are convenient, and to build a social learning community

Hew and Cheung (2014) identified several intrinsic motivators: desire to learn about a new topic or to extend current knowledge, curiosity, the desire for personal challenge and an extrinsic motivator: the desire to increase the number of certificates they earned. The findings regarding the registration phase are grouped and presented in Table 3.

Table 3: Structured findings regarding enrolment motivations

	Belanger & Thornton, 2013	**White et al., 2015**	**Hew & Cheung, 2014**	**Xiong et al., 2015**
Intrinsic motivations for enrolling				
For fun or entertainment	x			
MOOCs satisfy interest		x	x	x
Curiosity			x	
Challenge			x	
Extrinsic motivations for enrolling				
Recognition of accomplishment	x		x	x
Professional development	x	x		x
Supplement to credit bearing course			x	x
Other motivations for enrolling				
Interaction with students	x	x		
Convenient for people with limited access	x	x		
To experience online learning	x			

Activity phase

Studies that are grouped in the activity phase are concerned with the extent student engage and complete the course. There is a distinction made on "course level factors" and "student level factors" that influence student retention in MOOCs.

In her literature review in 2014, Jordan identified that most research so far had focused on looking at student's intentions and motivations to enrol, and she instead studied what characteristics of courses themselves predict course completion. She found no significant relationship between completion rate and the date the MOOC started, university rank or total number of students enrolled, but did find a negative correlation between completion rate and course length. This differed from the findings by Adamopoulos' (2013) study in which he did a sentiment

analysis of user reviews concerning 133 courses from 30 universities and 6 platforms. Adamopoulos also found a negative correlation with course length, but he did find a significant positive relation between university rank and completion rate. Other course level factors that had a negative effect were the difficulty, workload, and if the course was self-paced or not. Interestingly it was found that for difficult or longer courses, self-paced course design had a positive effect. Adamopoulos (2013) also found other positive effects when looking at the presence of final exams and project, open textbooks, peer assessment, and a certificate award for completion.

Hew (2016) found after studying three highly rated MOOCs, that there are five factors that engage students in MOOCs: course resources, instruction accessibility/passion, peer interaction, active learning, and problem oriented assignments with clear expositions. Course resources indicate that if there are variety of ways students are able to learn, it provides the availability for each student choose their favourite way and therewith engage. Active learning promotes application of new knowledge in any task or activity, instead of passively reading or listening. Problem oriented with clear expositions highlights that assignments should be oriented to solving real-world tasks. Unfortunately, Hew did not investigate the effect of these factors on course completion. Adamopolous (2013) found that the sentiment of students for assignments and course material has positive effects on the completion rate of the course and Gütl (2014) even mentions that poor course design was a reason not to complete a MOOC. Hone and El Said (2016) found that MOOC content has a significant effect on the perceived effectiveness of the course, which in turn significantly influences course retention. This al indicates that the MOOC content is also an important factor influencing retention rates, next to motivation and interaction with teachers and students. An overview of the course level factors findings is presented in Table 4, in which the relation with student retention is indicated as positive (pos.), negative (neg.) or neutral (neut.).

Table 4: Structured findings regarding course level factors effect on student retention

Course level factors	**Jordan, 2014**	**Adamopoulos, 2013**	**Gütl, 2014**	**Hew, 2016**	**Khalil & Ebner, 2014**	**Hone & El Said, 2016**
Self-paced (for longer courses)		Pos. / Neg.				
Course length	Neg.	Neg.				
Difficulty		Neg.				
Workload		Neg.				
Date of the MOOC	Neut.					
University rank	Neut.	Pos.				
Total number of student's enrolled	Neut.					
Hidden costs					Neg.	
Presence of final exams,		Pos.				
Presence of textbooks		Pos.				
Certificate or award		Pos.				
Overall course design			Pos.			
Diverse course resources				Pos.		
Active learning				Pos.		
Problem oriented with clear expositions				Pos.		
Sentiment for assignments and material		Pos.				
Perceived effectiveness of MOOC content						Pos.

Several authors looked at student level factors as predictors (Adamopoulos, 2013; Greene et al., 2015; Wilkowski et al., 2014). Greene et al. found that self-rated commitment was the biggest predictor, followed by prior experience, older students and more educated students. Wilkowski however, found no correlation between student's course completion and prior skill in a subject. They did find that students who completed course activities were more likely to earn certificates of completion, highlighting that it is important that students engage with the content. Lastly, Adamopoulos (2013) also found no correlation between the level of education or age of the students and completion rates.

Some studies in MOOC related literature examined student level factors by looking into what students themselves indicated their reason to engage in the course or to drop out was. In a literature study by Khalil and Ebner (2014) of 42 MOOCs, they identified lack of time, lack of motivation, feelings of isolation and lack of interactivity with students and professors, insufficient background and skills, and hidden costs. The lack of interactivity with the professors and students is also mentioned by several other authors (Adamopoulos, 2013; De Freitas et al., 2015; Hone & El Said, 2016), and comply with the findings of Hew (2016) in the previous section that both types of interaction engages students. Lack of motivation is also recognized by De Freitas et al. (2015), who state that students are unengaged and need to be motivated in order to increase completion rates. An overview of the student level factors is given in Table 5.

Table 5: Structured findings regarding student level factors effect on student retention

Student level factors	**Adamopoulos, 2013**	**Greene et al., 2015**	**Wilkowski & Russell, 2014**	**Hew, 2016**	**Khalil and Ebner, 2014**	**De Freitas et al., 2015**	**Hone and El Said, 2016**	**Xiong et al. 2015.**
Self-rated commitment		Pos.						
Activity completion			Pos.					
Prior experience		Pos.	Neut.		Pos.			
Age of students	Neut.	Pos.						
Level of education	Neut.							
Time available to student					Pos.			
Motivation (general)						Pos.		
Extrinsic motivation								Pos.
Intrinsic motivation								Pos.
Interaction with students	Pos.			Pos.	Pos.	Pos.	Pos.	
Interaction with teachers	Pos.			Pos.	Pos.	Pos.		

2.2.2 How can serious games impact student retention?

In the bulk of the literature that researched why MOOC retention is low and what factors influence it, different approaches are suggested to increase retention, some specifically mentioning serious games as possible method. De Freitas et al. (2015) state that a new form of gMOOCs (MOOCs that are gamified or use game-based approaches) are an interesting line of new research with great potential. Gütl (2014, p. 47) concludes his research on why students drop-out of MOOCs with the following:

"The next steps for future work are to include dynamic learning activities to capture the motivation and volition of the students to retain the students and the students' interests. This may include the use of serious games or gamification in the learning activities and ability to provide automatic feedback where relevant".

The fact that serious games are put forward as possible method to increase student retention, is probably based on key publications of Malone (1981), and Garris, Ahlers and Driskell (2002) who theorize that serious games are potentially highly motivating instructive environments, while at the same time several authors indicated that a lack of motivation was a significant factor causing low student retention (De Freitas et al., 2015; Khalil & Ebner, 2014).

In Malone's Toward a Theory of Intrinsically Motivation Instruction (1981) he theorizes that if students are intrinsically motivated to learn, they spend more time and effort on learning, feel better while doing it and apply their knowledge more in the future. Important factors that influence intrinsic motivation are challenge, fantasy and curiosity. In his later work together with Lepper, they add control to these three and group them into "individual motivations". They also add cooperation, competition and recognition as "intrapersonal motivation" factors (Malone & Lepper, 1987). Interestingly, these intrapersonal motivational factors could not only contribute to the intrinsic motivation of the game activity, but maybe also to decrease the lack of interactivity the students experience.

Garris, Ahlers and Driskell (2002) state that not only intrinsic motivation, but also extrinsic motivation is important to develop learners who are self-directed and self-motivated, because they then like the outcome and activity itself. While they mention it is important, they do not seem to account for extrinsic motivation in their proposed model. They propose an Input-Process-Outcome model of instructional games that is to be preferred over the traditional single-trial learning (e.g. read a book and be tested). The key component is the game cycle in which the users experience a repeated cycle of user-judgements, behaviour, and feedback that characterizes the engagement that players display. This engagement is described according to Csikszentmihalyi's (1990, p. 4) "state of flow, in which people are so involved in an activity that nothing else seems to matter". They propose that when instructional content and game characteristics are combined, this state of flow is triggered in the game cycle and an optimal state of performance is reached.

These two key publication form the basis for large amounts of serious gaming literature. The results of several meta-analyses on the impact of simulation, serious, or digital games showed that the game-based-instructional methods are indeed preferred over traditional teaching methods regarding learning outcomes (Clark et al., 2014; Sitzmann, 2011; Vogel et al., 2006; Wouters et al., 2013). However, there is no clear consensus on the motivational

impact. Sitzmann (2011) simply concludes that ironically the motivational effects of serious games are not considered in most studies, while this is what the effectiveness of serious games is based on in the first place. The meta-analysis Wouters (2013) found, when reviewing 31 studies, that serious games as teaching methods were not more motivating than conventional instruction methods. Wouters et al. give three possible explanations for this, the absence of autonomy of when to play and what to play undermines the intrinsic motivation, the favouring of the instructional design over the entertainment design resulting in an instructive but less engaging game, and the possible ineffectiveness of measuring motivation with post-course questionnaires. The meta-analysis conducted by Clark et al (2014) of 69 studies found positive results regarding motivational impact, but combined motivation with other aspects in their measurements, blurring the conclusion. It is important to note though that these meta-analyses hardly contain comparisons of serious games and traditional teaching methods in online environments, and none in a MOOC environment.

Next to their motivational appeal, serious games are considered by Hew (2014) as a possible method for increasing user engagement as they are a form of active learning. Interestingly, serious games could also influence maybe two out of the four other factors that influence student engagement in MOOCs according to his research. Serious games could be used to apply knowledge in assignments with clear expositions, and increase the number of course resources a student has at his or her disposal.

2.2.3 What are findings of serious game related MOOCs studies?

Despite the growing bulk of literature surrounding MOOCs since they first appeared in 2008 (Liyanagunawardena et al., 2013), only a total of 9 papers were found by 7 different authors (or author groups). Interestingly, all papers were published in conference proceedings ranging from 2013 to 2017. The three perspectives the papers show are a *technical perspective* focused on the integration of a serious into a MOOC, a *course perspective* focused on changing the whole course into a serious game and the *supportive perspective* focused on using serious games as supportive part of the course. All three perspectives give valuable insights and are discussed in this section.

Technical perspective

Three of the papers discussed the progress made with integrating a serious game into edX (Blanco, Serrano, Martinez, & Fernandez-Manjon, 2013; Freire et al., 2014; Freire, Martínez-Ortiz, Moreno-Ger, & Fernández-Manjón, 2015). In their most recent work, Freire et al. (2015) have analysed the integration of serious games into two open source MOOC platforms: edX and OpenMOOC. They evaluate different integration options in edX and identified two major decisions that facilitators have to make for game-platform integration, and describe a catalogue of best practices from technical and educational perspectives on the game, course and platform design (Table 6). While the best practices on game and course level can be taken into account for the design of the serious game in this study, the best practices for the platform are less useful, as it is not possible to change the course itself.

Table 6: Best practices in SG integration in MOOC (Freire et al., 2015)

Component dimension	Technical	Educational
Game	- Support standardized packaging formats, providing search-friendly metadata. - Simplify integration into host courses and platforms.	-Use student profiles to customize learning experience. -Provide information on student actions and progress to the MOOC platform. -Allow authoring access to the educational content, so that course creators can tweak it.
Course	-Integrate guest activity into general course progression provide activity with student profile, -Reflect in-game decisions in updated profile.	-Provide educational context for SG, both before and after. -Select the SG activities that best advance the educational goals of the course.
Platform	-Clear APIs for guest activities. -Support several (incremental) integration levels.	-Collect and facilitate analysis of data from guests, allowing, for example, student leaderboards to be generated. -Support A/B testing of game variants.

Course perspective

Two other studies describe an approach to transform an entire MOOC into a serious game. In the article by Tan (2013), the proceedings of the conceptualization of a "MOOC Game" is described, and the first steps of design focused on the challenge and play aspects of the game are mentioned. Maalej et al. (2015) transformed the MOOC environment into a virtual learning environment in which a student creates an avatar and is able to choose their own sequence, or path, of the course activities by walking around their avatar, selecting activities and interacting with avatars from other students. Unfortunately, they did not experiment with the game yet and no effects on student retention are researched.

Supportive perspective

Romero and Usart (2013) incorporated two different serious games in a MOOC in which students could orient on their level of entrepreneurship. The students had to complete four main topics, and in the last two topics the students could play a serious game. After a qualitative analysis the authors found that students highly appreciated the presence of the two serious games, and they state that using serious games could possible contribute to decrease the low drop-out rates experienced by MOOCs. However, looking at the different goal of this MOOC (orienting instead of learning), the low number of participants (45) and the focus of teachers to interact with students as much as possible, make that the findings of this research are not considered generalizable to other MOOCs and serious game integrations, as MOOCs are known for having inherently low student-teacher interaction due to their massive nature (Hone & El Said, 2016).

Vaibhav and Gupta (2014) did not research the integration of a serious game in a MOOC, but instead studied both separately and combined the results. First, they do a region and discipline analysis of MOOCs to find some general descriptive statistics. Then they proceed to compare a group of students who use traditional methods to learn a list of words, with a group that use a "gamified method" in the form of a quiz. They call this a form of gamification, but according to the definition in this research the use of a quiz is considered a serious game, since it is not a game element but a game itself. Both groups are asked to fill in a test after 3 days, and they conclude that from the game-group more students showed up, passed the test and had more fun doing it. However, they then seem to "just combine" this with the MOOC descriptive findings and conclude that if MOOCs are gamified, higher student engagement can be obtained, which is not deemed a valid way of reaching this conclusion.

Bernaert, Crepon, and Dhorne (2015) made a serious game that covered the learning content and offered it as main instruction method in the MOOC and as evaluation method at the end of each week. The MOOC had a completion rate of 31% for enrolled students, which is high considering the average completion rate of 6.5% (Jordan, 2014). The authors note that 52% of the students thought the serious game helped in their understanding of the course, and 43% think it increased their interest in the course. 38% of the students indicated that they found the serious game the most enjoyable part of the course. When this group was compared with the group that indicated they found another course component the most enjoyable, they found that no significant differences exist between professionals and students, and for different age groups, but women were more present in the game-group. However, the contribution of the serious game to the high completion rate is debatable, since the data cannot be compared to the same MOOC without a serious game and thus other factors could have influenced the completion rate.

2.3 Theoretical framework

The theoretical framework of this research was developed based on the findings in the literature study in Section 2.2. First, it is explained which factors were selected for the theoretical framework and why in Sub-section 2.3.1. Then, in Sub-section 2.3.2 the selected factors are used to develop a theoretical framework consisting of three parts, each representing one of the sub-research questions 3, 4, and 5 shown below. To be able to answer these research questions, in Sub-section 2.3.3 clear and measurable hypotheses are formulated that can be tested with empirical research.

3. *To what extent do extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality influence student retention?*
4. *To what extent does a serious game affect extrinsic motivation, intrinsic motivation, interactivity, and perceived MOOC quality?*
5. *Is there a difference in student retention when a serious game is incorporated into a MOOC?*

2.3.1 Selecting factors for theoretical framework

The theoretical framework is developed by identifying and structuring factors that have been found to affect student retention in previous, and that might be influenced by serious games.

In the literature study on the registration phase of MOOCs in sub-section 2.2.1, it was found that different student types exist in MOOCs, and that students have different intentions and reasons for enrolling (see Table 2, page 29). The student types are not considered relevant for this theoretical framework as they are not factors influencing student retention. Student intentions for enrolling is a factor influencing student retention, but is not accounted for in this theoretical framework as changing the intention of students to complete resembles more the goal of the entire framework, than an individual factor that could be influenced. Instead, a more detailed look was given into the reasons student enrol in a MOOC.

It appears students enrolled for extrinsic reasons, intrinsic reasons, to interact with other students, because it was convenient and to experience online learning (see Table 3, page 30). The latter two enrolment motivations were dismissed to be part of the theoretical framework as it was presumed these could not be influenced by a serious game during the course. This left the intrinsic motivation, extrinsic motivation and social motivation to enrol as possible candidates for the theoretical framework inclusion.

Intrinsic motivation of students was found an eligible factor for the theoretical framework as in Sub-section 2.2.2 it was shown that serious games are known for their intrinsic motivational capabilities (Csikszentmihalyi, 1990; Garris et al., 2002; Thomas W. Malone, 1981) and Xiong (2015) found using a Structural Equation Modelling (SEM) approach that both intrinsic and extrinsic motivations to enrol were a significant predictor for student engagement in a course.

As *extrinsic motivation* seems to be an important reason for students to enrol in a MOOC and Xiong (2015) found that extrinsic motivation was a predictor for student retention, this indicates that extrinsic motivation is an important factor influencing student retention in MOOCs. However, too our knowledge, it has not been researched if serious games could be designed to increase extrinsic motivation. Due to the presumed importance of extrinsic motivation on student retention, this factor is incorporated into the framework, noting that this factor had to be purposefully accounted for in the serious game design.

Interactivity was incorporated into the theoretical framework as interaction with other students and teachers is not only a much mentioned reason for students to enrol in a course (Belanger & Thornton, 2007; White et al., 2015), but the lack of interaction is also a much mentioned reason of students to drop out of a MOOC (De Freitas et al., 2015; Khalil & Ebner, 2014), and several authors found a significant correlation between interactivity and student retention (Adamopoulos, 2013; Hone & El Said, 2016). Besides the relation with student retention, it is also specifically mentioned by Malone and Lepper (1987) that interaction with other students is an important way games achieve their intrinsic motivational learning environments, showing the possibility games offer to have students interact with other students.

In the literature study on the activity phase in Sub-section 2.2.1, several course level factors were found to affect student retention in some way (see Table 4, page 31). Neutral course level factors like the starting date of the MOOC and the number of enrolled students were not found to be eligible for the theoretical framework. Furthermore, the course level factors that cannot be influenced or changed with a serious game were excluded. Examples of these are the course length, if the course is self-paced, the university rank, hidden costs, presence of final exams, presence of textbooks and a certificate or award.

The difficulty and workload of the MOOC might be influenced by incorporating a serious game, depending on the type and complexity of the game. A serious game presence does increase the diversity of course resources, and could also enable active learning by applying knowledge in the game, could be used to clarify problems with clear expositions, influence the sentiment for assignments and material, influence the perception on overall course design and the perceived effectiveness of the MOOC. These factors were grouped under one “umbrella term”: *perceived MOOC quality*. It is believed these factors can be grouped as they could all be interpreted as factors that influence how students rate the quality of the course. This is more apparent for factors like the grade for overall course design, perceived effectiveness of the MOOC, and sentiment for assignments and material factors, but less clear with diversity of course resources, active learning and problems with clear expositions factors. The assumption these factors can be grouped into one term had to be taken into account when discussing the findings.

To conclude, the factors that influence student retention and might be influenced by incorporating a serious game and are therefore included in the theoretical framework are:

- *Extrinsic motivation*
- *Intrinsic motivation*
- *Interactivity*
- *Perceived MOOC quality.*

2.3.2 Theoretical framework description

In the previous section four factors are described that influence student retention in MOOCs, that might be influenced by incorporating a serious game, therewith providing the theoretical basis that the incorporation of a serious game into a MOOC could lead to improved student retention. These three relations are depicted in the high level visualisation of the theoretical framework in Figure 4. First it is explained why the theoretical framework consists of three separate parts, before each part is discussed in more detail.

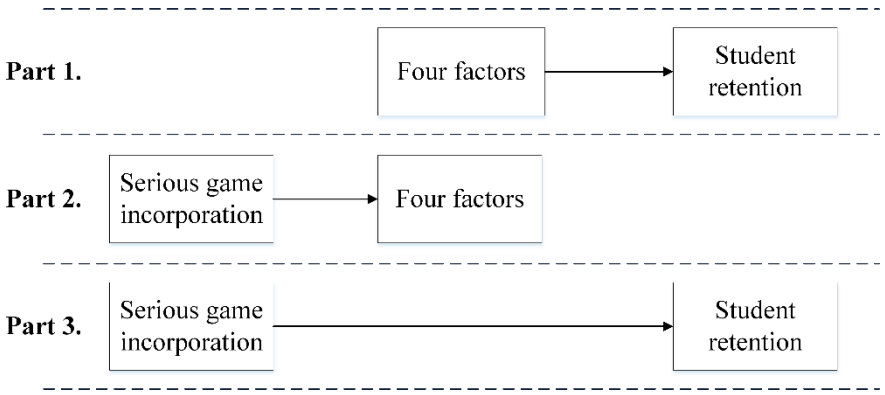


Figure 4: High level visualisation of theoretical framework

It was chosen to have the theoretical framework separated into three parts to reduce the research complexity. To develop a one model theoretical framework, an in depth pedagogical literature study would have been required to determine if the factors are moderator or mediating variables, and would have to be tested with more complex statistical methods. Such a model would provide better insights in the relations between the factors, but was only considered worth the increased research complexity if it could be stated with more certainty that the four factors could be influenced by a serious game. The lack of knowledge on these relations is the reason this study was conducted in the first place, which is why it was decided to first analyse the three aspects separately, to see which relation is worth focussing future research efforts on with a more complex research approach. That being said, this theoretical framework is still considered to be able to give insights in how an effect of the serious game on student retention might be explained, by combining and discussing the findings of each of the three parts of the theoretical framework.

Part 1: The factors influencing student retention

The first part of the research is concerned with the relation between the four factors with student retention, but also with the four factors amongst themselves. This theoretical framework gives insight into to what extent the identified factors could influence student retention. This means the four factors are the independent variables, and the student retention variables are the dependent variables.

The relation between the four factors and student retention has a relative strong theoretical and empirical basis as shown in Sub-section 2.2.1, with several studies indicating their effect student retention. However, only two of those studies were found to have researched relations amongst these factors in MOOC environments. Hone and El Said (2016) used a Partial Least Squares to perform a structural model analysis and found no significant path-

coefficient between interaction and the perceived effectiveness of a MOOC (which in this study was grouped under perceived MOOC quality). Xiong et al. (2015) used a Structural Equation Modelling approach to research relations amongst social motivation, extrinsic motivation, intrinsic motivation, engagement and retention. They found significant relations between the three types of motivations, where the social motivation is measured with the extent to which students signed up to connect with other students, or because they have friends taking this course. The social motivations therewith somewhat resemble the interactivity factor in this study. These two studies do not provide an empirical basis to specifically include or exclude relations amongst the four factors, which is why these relations are all incorporated into the first part of the framework, meaning this part of the theoretical framework is also used to explore to what extent the factors are related. A detailed visualisation of the first part of the theoretical framework is presented in Figure 5.

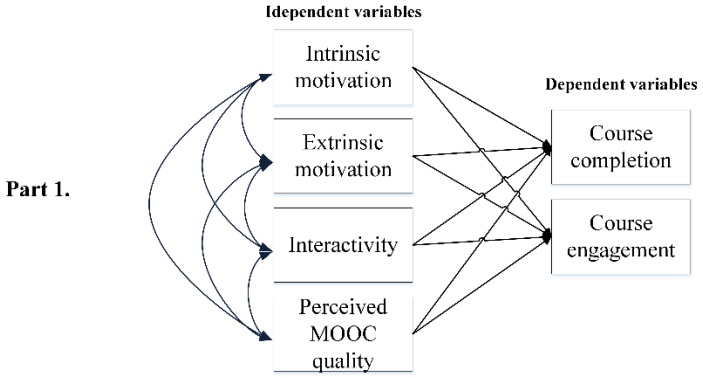


Figure 5: Detailed visualisation of part 1 of the theoretical framework

Part 2: The effect of the serious game incorporation on the four factors

The second part of the theoretical framework is concerned with the effect of the serious game incorporation on the four factors. This means in the case the serious game is the independent variable and the four factors are the dependent variables. The relations between the four factors are not accounted for, as these are already researched in the first part of the theoretical framework. A detailed visualisation of the second part of the theoretical framework is presented in Figure 6.

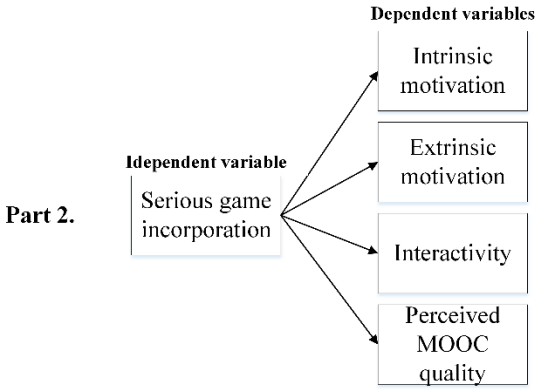


Figure 6: Detailed visualisation of part 2 of the theoretical framework

Part 3: The effect of the serious game incorporation on student retention

As stated in the introduction of this section, the first two parts of the theoretical framework provide the basis for a relation between the serious game incorporation and an effect on student retention. The third part of the research is concerned with if a difference in student retention can be observed by incorporating a serious game, and therewith prevent not measuring a difference in student retention because the serious game affects student retention in another way than via the four identified factors. In this case the serious game incorporation is the independent variable, and student retention is the dependent variable. A detailed visualisation of the third part of the theoretical framework is presented in Figure 8.

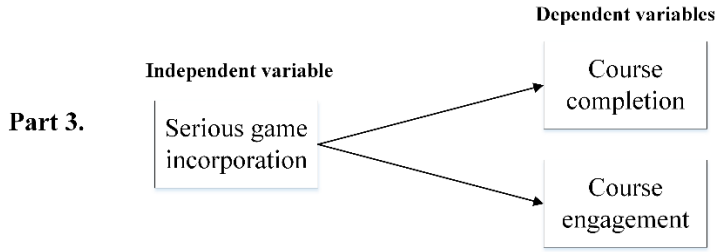


Figure 7: Detailed visualisation of part 3 of the theoretical framework

2.3.3 Hypotheses formulation

With the development of the theoretical framework the main research question is not answered as the explanatory nature is not addressed. The theoretical framework only shows the theoretical extent of how a serious game could affect student retention in MOOCs. By formulating hypotheses for these relations, the theoretical framework can be evaluated and a more empirically based answer could be given to the main reason question: *To what extent can student retention in MOOCs be improved by incorporating a serious game?* In order to construct clear hypotheses, the term student retention must be operationalized first.

Operationalization of "student retention"

The definition of student retention by Berge and Huang (2004, p. 3), "retention is the continued student participation in a learning event to completion, which in higher education could be a course, program, institution, or system", was chosen as it does not solely focus on completions rates, but also on engagement throughout the course.

In the available literature there are several ways to measure completion rates. In her literature review Jordan (2014) found that earning a certificate was the most used one. Others were completing all the assignments, achieving a strong final score, or taking an exam. In this research a student is considered a completer if he or she obtained a sufficient grade for the course.

The hypotheses concerning the course engagement is split up into the effect of the factors on the number of forum posts students made, the number of assignments students made, and the number of videos students watched to measure the engagement, as several authors have used these to indicate course engagement (Hew, 2016; Khalil & Ebner, 2014; Kizilcec et al., 2013; Xiong et al., 2015).

Hypotheses

The hypotheses regarding the factors and their influence on student retention are shown in Table 7. As there are both empirical and theoretical findings backing up these relations, the hypotheses are formulated to have a positive effect on student retention. There are no hypotheses drafted for the relations amongst the factors, as these are of exploratory nature.

Table 7: Hypotheses for evaluation of theoretical framework part 1

1. Intrinsic motivation	
H1.1.1	Students who completed the course have a higher intrinsic motivation than students who did not complete the course.
H1.1.2	Intrinsic motivation has a positive impact on the number of forum posts
H1.1.3	Intrinsic motivation has a positive impact on the number of assignments made per student
H1.1.4	Intrinsic motivation has a positive impact on the number of videos watched
2. Extrinsic motivation	
H1.2.1	Students who completed the course have a higher extrinsic motivation than students who did not complete the course.
H1.2.2	Extrinsic motivation has a positive impact on the number of forum posts
H1.2.3	Extrinsic motivation has a positive impact on the number of assignments made per student
H1.2.4	Extrinsic motivation has a positive impact on the number of videos watched
3. Interactivity	
H1.3.1	Students who completed the course have experienced more interaction than students who did not complete the course.
H1.3.2	Interactivity has a positive impact on the number of forum posts
H1.3.3	Interactivity has a positive impact on the number of assignments made per student
H1.3.4	Interactivity has a positive impact on the number of videos watched
4. Perceived MOOC Quality	
H1.4.1	Students who completed the course have a higher perceived MOOC quality than students who did not complete the course.
H1.4.2	Higher perceived MOOC quality leads to higher forum activity
H1.4.3	Higher perceived MOOC quality leads to more assignments made per student
H1.4.4	Higher perceived MOOC quality leads to more videos watched

The hypotheses regarding the factors and their influence on student retention are shown in Table 8. In case of intrinsic motivation and interactivity there is a theoretical basis that serious games can influence these factors positively. However, the effect on extrinsic motivation and perceived MOOC quality lack this basis, which is why they are formulated with no specific direction.

Table 8: Hypotheses for evaluation of theoretical framework part 2

H2.1	The serious game has a positive impact on intrinsic motivation
H2.2	The serious game has an impact on extrinsic motivation
H2.3	The serious game has a positive impact on interactivity
H2.4	The serious game has an impact on perceived MOOC quality

The hypotheses regarding the factors and their influence on student retention are shown in Table 9. As not all of the relations in the first and second part of the theoretical framework are positive, and there is little empirical findings of other studies available with this being one of the first studies to incorporate a serious game into a MOOC, the hypotheses are formulated with no specific direction.

Table 9: Hypotheses for evaluation of theoretical framework part 3

H3.1	The serious game has a positive impact on course completion
H3.2	The serious game has an impact on number of forum posts
H3.3	The serious game has a positive on number of assignments made
H3.4	The serious game has an impact on number of videos watched

2.4 Chapter summary and conclusions

The aim of this chapter was to develop a theoretical framework represented with hypotheses that can be tested with an experiment, by answering the first two research questions:

1. *What factors influence student retention in MOOCs?*
2. *Which factors influencing student retention could be affected by a serious game?*

Four factors were identified that influence student retention in MOOCs and might possibly be influenced by incorporating a serious game:

1. Intrinsic motivation
2. Extrinsic motivation
3. Interactivity
4. Perceived MOOC quality

With this, research questions 1 and 2 are answered. The factors and their theoretical relations with student retention and serious games were visualized in a theoretical framework, consisting of three parts; (1) the relation between the factors and student retention, (2) the relation between the serious game and the factors, and (3) the relation between the serious game and the student retention. For every presumed causal relation a hypothesis was drafted. Testing these hypotheses leads to the explanatory part of this research: the experiment in which all three parts of the theoretical framework are evaluated.

3. Experimenting with a serious game in a MOOC

To test the hypothesized influence of the serious games on the student retention in a MOOC, an experiment was conducted. This experiment took place in the Create Problem Solving and Decision Making (CSPD) MOOC offered by the TU Delft. First the MOOC is analysed and described in Section 3.1. This MOOC description and the theoretical framework provided the input for the used serious game, described in Section 3.2. Lastly the experimental set up is described in Section 3.3. In the experimental set up several analyses are conducted that were required to prepare the analysis regarding the three theoretical framework parts, which is why these analyses are part of this chapter.

3.1 The CPSD MOOC description

The experiment is conducted in the Creative Problem Solving and Decision Making (CSPD) MOOC. The CSPD MOOC is hosted on the edX platform, and teaches an analytical approach to solve complex problems. At the start of each course the students choose a case concerning a complex problem, and apply several analytical methods to be able to show the owner of the problem what his alternative courses of action are. The MOOC requires no payment to enrol but the students are able to pay beforehand, during the course, or after completing it to earn a verified certificate from DelftX, but only if they complete the course with a sufficient grade. The MOOC is self-paced and takes approximately 20 to 40 hours to complete. After an introduction step that explains the course and the edX platform, the students are asked to fill in the pre-questionnaire. Then, the students are guided through five different steps, in which one or several methods are learned. Figure 8 shows what one step looks like. At the end of the course there is a brief video to wrap up and a post-questionnaire to evaluate students' experiences.

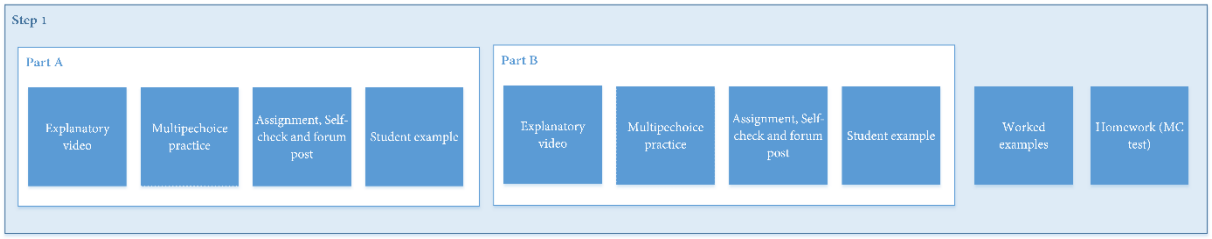


Figure 8: A visialisation of one step in the MOOC

Each step of the course consists of one or two "parts". Each part consists of the same "components"; an explanatory video that explains the method, some multiple choice questions to practice knowledge from the video, the assignment in which the students have to apply the method to their case, followed by the peer-assessment of another students' assignment, and a student example of the assignment chosen by the instructors. If the part(s) is done, the students are shown a worked example that shows the (combined) application of the method(s) on an illustrative case by the instructor. Next, the different steps are briefly discussed before an overview is presented in Table 10. The steps and overview are explicitly stated as they serve as input for the serious game design.

Step 1: Actor Analysis. In this step the students are taught that in order to solve a problem, you must understand it first. The first step for doing this is identifying what actors (people, groups, and organizations) are involved and how they perceive the problem. For this the students are learned how to perform an *actor analysis* (part 1) and how to correctly formulate a *problem statement* (part 2) for each relevant actor.

Step 2: Goal Analysis: The second step builds on the first step by analysing what the goal of each of the actors is and with what criteria they can be measured. This is done to make the problem more explicit as it is usually hard to measure high level goals of actors. To operationalize the goals of the actors into criteria the *goal tree* method is used. This step only contains one part.

Step 3: Causal analysis. When the goals of the actors are operationalized into measurable criteria in step 3, it is possible to make a *causal relations diagram*. In this diagram causal relations are mapped that affect the criteria. This is done to be able to understand the current problem situation better and later on see how these factors can be influenced with possible solutions. Just like step 2, this step only contains one part.

Step 4: Alternatives analysis. This step contains two parts. In the first part the findings from the first three steps are combined into a *problem diagram*. This overview shows alternatives that influence the criteria via the causal relations, and can therefore be used to get a quick overview of the problem situation. Furthermore, this problem diagram allows to evaluate how possible future *scenarios* could influence the effect of the alternatives on the criteria. Students are made to think about possible relevant future trends or changes.

Step 5: Decision analysis. In the final step the students' are taught methods that help with making the decision on what solution to implement. In part one the effect of the alternative on the criteria is *estimated* using several techniques. This is done as finding undisputed data with these types of problems is often difficult. In part two the different alternatives are compared using *Score ards* and *SMART* (Simple Multi Attribute Rating) *tables.* In the score card the estimated scores of the alternatives on the criteria are presented. The SMART table is then made for each individual actor by adding a weight to each criterion that corresponds with how important that specific actor finds that specific criterion.

Table 10: The CSPD MOOC steps, methods and lessons

	Methods	**Lessons**
Step 1	-Actor analysis	Who are involved?
	-Problem statements	How do they perceive the problem?
Step 2	-Goal tree and criteria	What are their goals, and with what criteria can we measure them?
Step 3	-Causal diagram	What currently influences these criteria?
Step 4	-Problem diagram and alternatives	What are solutions that positively influence the criteria?
	-Construct scenarios	How could the current situation change?
Step 5	-Estimates	What do you guess the effect of the solutions would be?
	-Score care / SMART table	How do alternatives score on the criteria in each scenario?

3.2 The serious game description

This section describes the serious game that was used in the experiment. This is done by describing the requirements and priorities the serious game had to comply with (Sub-section 3.2.1), the serious game design (Sub-section 3.2.3) and how the game was integrated into the MOOC (Sub-section 3.3.3).

3.2.1 Design space

The design space is the imaginary intellectual space that contains or envelops all of the potential solutions (Dym, Little, Orwin, & Spjut, 2009). The design space of the MOOC is bounded by constraints and objectives. In this research, the design space is depicted with the MoSCoW method, a method that can be used help understand and manage priorities (Dynamic Systems Development Method ltd, 2014). An explanation of the MoSCoW method can be found in Appendix 1. The MoSCoW table showing the design space for this serious game is presented in Table 11.

When using the InThere game design approach, the requirements and priorities (often the problem to be solved) from the client's point of view are identified in a so-called Gamestorm. The Gamestorm is a workshop in which the game designer and the client together brainstorm on the problem to be solved, the goal to be achieved and the game to be used. In this study, in preparation for the Gamestorm, the CSPD MOOC facilitated in 2016 was analysed and meetings were held with the TU Delft Extension School. From the analysis and the meetings it can be concluded that most students were working professionals that seek to improve their career or want to at least apply it in their daily work. This is taken into account in the serious game design by ensuring the game does not come across as childish and give the participants the feeling they are not taken seriously. Furthermore, it is important to emphasize with the game that the course is indeed applicable to their daily life. From the qualitative results it can be concluded that the general approach of the course is highly valued and the methods are often specifically mentioned, albeit in a negative or positive way. To emphasise this course's apparent strong point, the methods and the general approach should be prominent in the game design. A detailed description of the CSPD MOOC 2016 analysis can be found in Appendix 2.

Table 11: MoSCoW prioritizing of the "serious game MOOC"

The game...	
Must have…	…at least one game incorporated into the MOOC
	…a clear use for people of different backgrounds
	…a possibility for all enrolled students in the game-group to access the game
	…24/7 availability due to the self-paced nature of the MOOC
	…the capacity for a large group of people (500) to play the game
	…a high reliability regarding user data streams
	…a development time under 8 weeks
	…a budget requirement under €10.000
	…a high security level to prevent student data leakage
	…a positive effect on the motivation, interaction and MOOC quality perception.
Should have…	…an integration of the game into edX (not downloading the game and play locally)
	…no access for people from the control-group
	…the capacity to track user data (e.g. scores, behaviour)
	…a single single-in system to lower the barrier for playing the game
	…a relation with the course content
	…do not decrease the previous mentioned factors either or lower his/her scores
	…similarity to another of InThere's games to minimize the development time and costs.
	…a professional appearance and not give a childish vibe
	…a leaderboard with individual user statistics
Could have…	…an overarching storyline that connects the games
	…an access restriction to people that did not finish the corresponding steps yet
	…a content representation that is interpreted the same by all students
Won't have…	…an interactive deployment of the storyline based on the student's results in the game and in the course
	…an individual texts and storyline according to the student's enrolment motivation, intention, and preferred way of learning as indicated at the start of the course.
	…Students directly interacting with other students inside the game environment

3.2.2 Game design

The serious game consists of 13 components; five games and eight multiple choice question replacements. Both components types and their interaction are described in this section to give a complete picture of the serious game intervention. The five games differ as they resemble the content of the step they belong to (for step content, see Section 3.1). Each game consists of several rounds in which the student has to choose the right answer by clicking on something or dragging something to the right place. Each round has a limited amount of time. The games are individually elaborated upon further, after the relation with the multiple choice questions is explained.

The new multiple choice questions contain the exact same questions and answers as the original multiple choice practice questions, but are changed visually to correspond with the game appearance. The student encounters the questions right after the explanatory video, just like in the original course. Before they start the questions a message

is shown that explains that these practice questions are not of influence on their grade, but they do give the opportunity to earn a *multiplier* for the game at the end of the step. This multiplier increases the score that is achieved with the game. For each question they answer right, their score increases by 2%. So, if there are three practice questions, they can earn a multiplier of 1.06 x game score. In case the step consists of another multiple choice practice component, the multiplier could increase to 1.12 x game score. Naturally, this means that an individual user-ID of the student is required that remains the same throughout the course. The students can only play the practice-multiple choice questions once, to prevent them from trying to get a higher multiplier "now that they know the answers". To still ensure the students can learn from the multiple choice practice the questions, the right and wrong answers and an explanation per question are shown the second time they reach the webpage containing the game. The relation between the multiple choice questions and the game are illustrated in Figure 9.

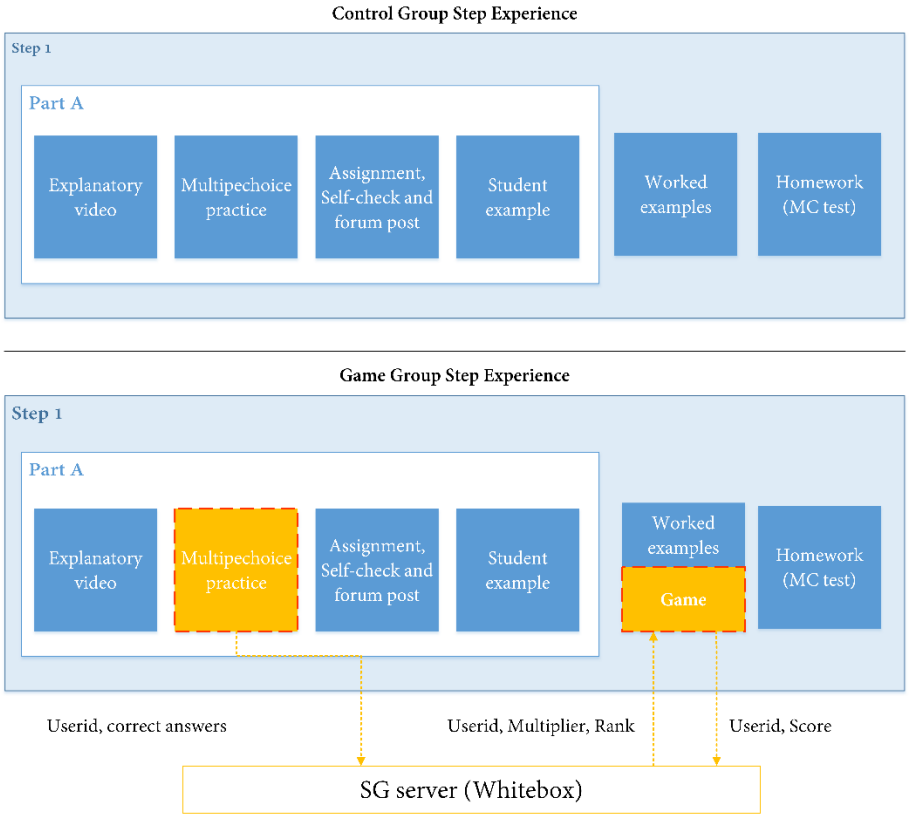


Figure 9: Relation between multiple choice components and game in a step

When the student arrives at the game at the end of the step, they are first shown an introduction screen. This screen contains a text message, an illustration of the game to come, and a start button. In the text message the students are first welcomed and reminded that they are trying to become a "creative problem solver". This reminder is explicitly made to make students remember why they are following the course, and therewith trigger their enrolment motivations. As different students have different motivations for enrolling, no explicit motivator (e.g. relevance for job) is mentioned and the reference is kept vague. The second part of the message briefly explains the game at hand by stating the goal and how this goal is achieved. Lastly, some textual encouragement is used to make the students realize the game is also "for fun", meaning to trigger their *intrinsic motivators*. The intro screen for game 1 is shown in Figure 10.

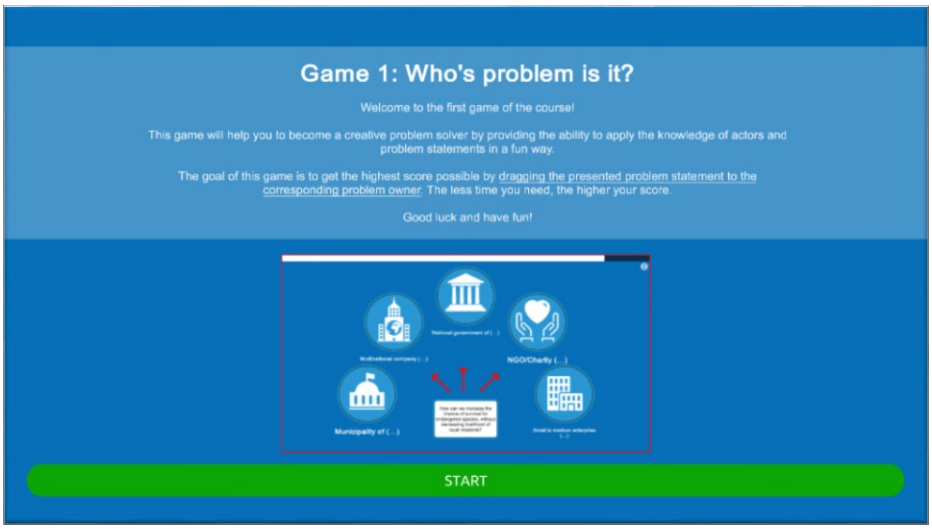


Figure 10: Introduction screen to game 1

When the game is finished, two other screens are presented (Figure 11). The first screen gives a quick overview of the score they obtained, the multiplier they earned from the practice multiple choice questions in this step, and what their final score amounts to. This final score is then compared to other players' highest scores, and the current rank they player holds is presented. This ranking system in important to give student the feeling of competition with other students, therewith possibly improving their *student to student interactivity*.

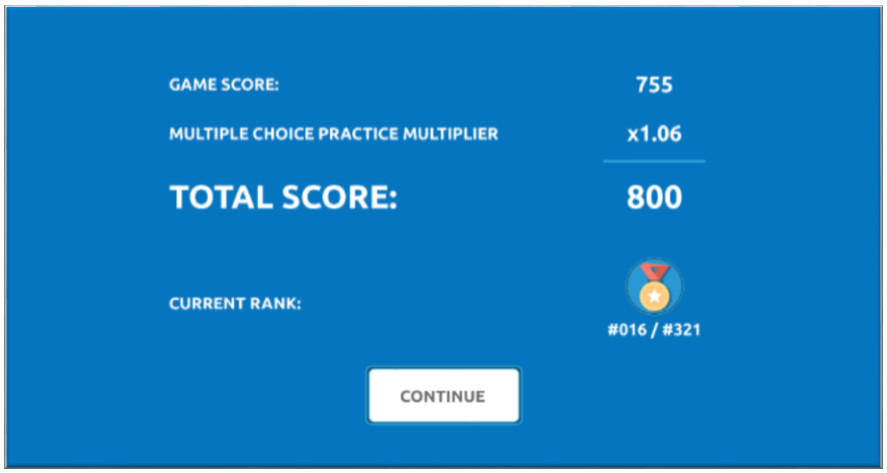


Figure 11: First screen after a game

The second outro screen (Figure 12) encourages the student to play the game again and get a higher score. To ensure learners reflect on what they are doing and learn from it, as well as to prevent students from being frustrated if they keep getting low scores, common pitfalls are shown. After that it is emphasized that the game does represent the content and can be seen as an indicator of their knowledge, encouraging students to study more and to prevent students from skipping the games. Students are also encouraged to discuss their game experience on the discussion forum, to increase their *interactivity*. A short reference to the next game is given to trigger their curiosity, which is an *intrinsic motivator*.

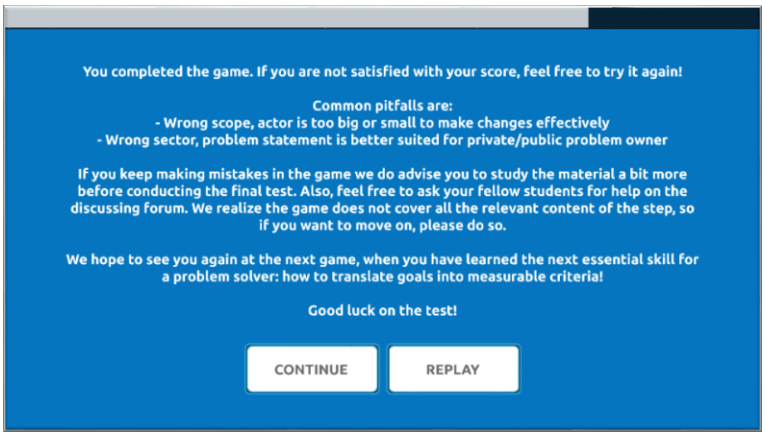


Figure 12: Second screen after a game

Game 1: Whose problem is it?

The first game (Figure 13) corresponds with the first step. The player is supposed to drag the problem statement (in white) to the corresponding problem owner. The actors the student can choose from are deliberately similar to one another, to make students realize how small adjustments in problem formulation can make a difference for the demarcation of the problem. When the students has given his or her answer, the problem statement disappears and a new problem statement is shown. At the bottom of the screen the student can see how many questions are left by counting the circles. The circle that is blue indicates which challenge the student is currently at, while the red or green circles indicate if the previous answers were right or wrong.

Figure 13: Game 1: Whose problem is it?

Game 2: Guessing Goals

The second game (Figure 14) concerns the goal-tree method that was taught to the students in the second step. Before the time is over, the students have to drag one of the three white goals to the darkened area. The students have to figure out which goal is the correct one by analysing the lower goals. This makes them think about how goals are supposed to be operationalized in a goal tree. Furthermore, sometimes two options on the left appear to be correct, but is one formulated incorrectly. This is done to make students think about common pitfalls when constructing a goal tree for their case.

Figure 14: Game 2: Guessing Goals

Game 3: Capturing Causalities

The third game (Figure 15) is about the causal analysis method used to map causal relations amongst relevant factors. One of the pitfalls in this step is wrong interpretation of the relation between arrows. When a "-" is indicated at an arrow, this is often mistaken for a "regular" negative effect, while in reality it means: "If factor A decreases, B increases and vice versa". This common pitfall is translated in the game by having the students choose between the + and – sign. Because this game is quite straight forward, the signs and the factors change places and the time per challenge was shortened.

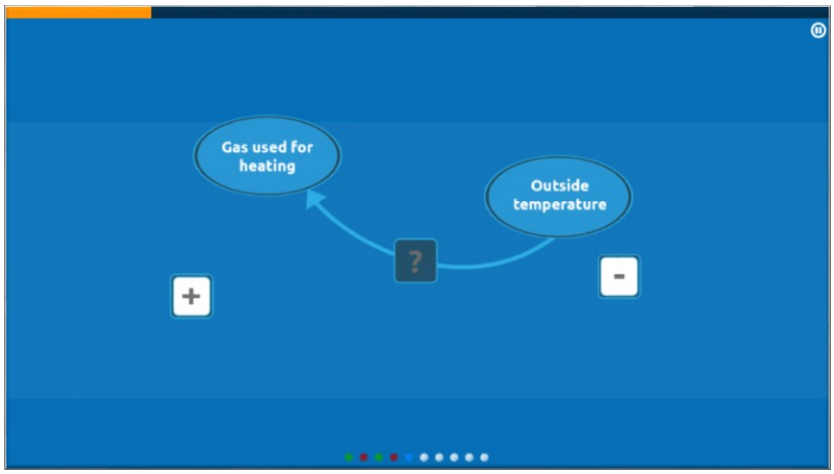


Figure 15: Game 3: Capturing Causalities

Game 4: Alternative Influences

The fourth game (Figure 166) is a partial representation of the problem diagram used in step 4 to structure the problem situation. It also incorporates that the students have to look at the influence of an alternative at the problem situation, and what happens if the effect of the alternative increases or decreases in the future. The students have to look at the arrow above or below the alternative, follow the causal relations and indicate if the effect on the criteria increases or decreases.

Figure 16: Game 4: Alternative Influences

Game 5: SMART Calculating

The final game (Figure 17) tests the student's understanding of how to read a SMART table to determine which alternative scores highest for the actor for whom the weights were assigned. Because often a lot of SMART tables need to be made (one for each actor), quickly being able to analyse them could come in handy. However, because this "number crunching" is not necessarily fun to do for a long time, only 5 challenges per game have to be done before students see their score.

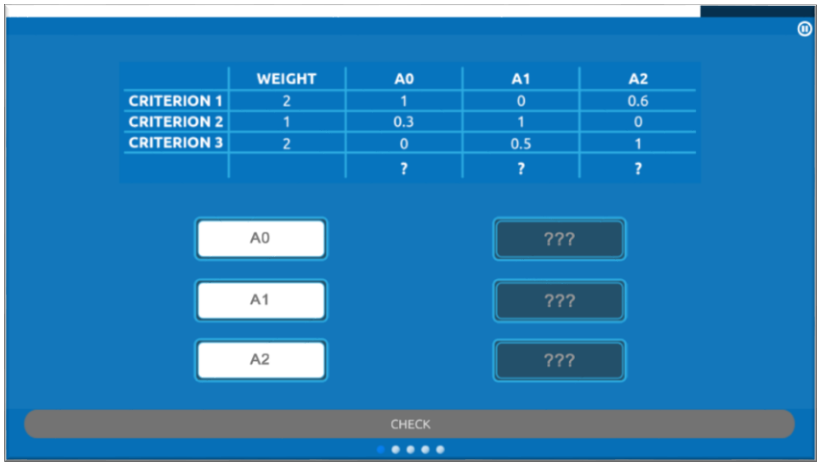


Figure 17: Game 5: SMART Calculating

3.2.3 Game integration

Before the serious game design is presented the integration of the game into the MOOC is discussed, as this has a large impact on the possible design choices. Freire et al. (2015) describe that there are two major decisions concerning game-platform integration. A stand-alone game or an in-browser game, and an integrated serious game server or an external serious game servers. The choices are illustrated in Figure 18.

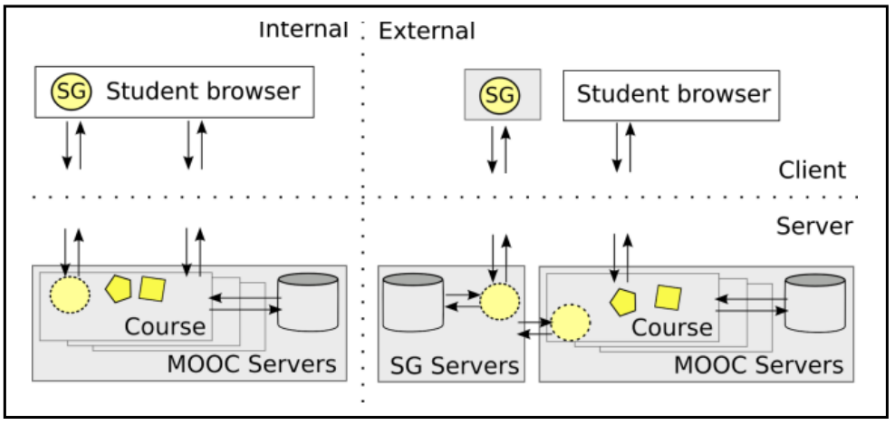


Figure 18: SG-platform integration decisions (Freire et al., 2015)

For this study an in-browser game is used. This prevents that students have to go through installation procedures, which raise the barrier to play the game. It also reduces the development time as the game does not need to be developed for mobile phones, tablets or different operating systems. This does require the students to follow this particular part of the course on their computer. However, the collective data of all DelftX MOOCs showed that relatively few students follow MOOCs via other devices than computers. To make the small group of students who do uses other devices aware of this specification, a message in the course was written asking them to follow this part on their computer via a web-browser.

On the server side the choice has been made to host the game on an external server. The reason for this was that access to the players data was required and this could not be guaranteed when the game was integrated in the edX server. Freire et al. (2015) point out that hosting a game could be taxing on the servers, but fortunately the servers that are used in this research were specifically designed to process large amounts of game data. Having an external server does pose the difficulty of how the serious game will be integrated into edX. Freire et al. (2015) list three possibilities for this, but state that all three requires considerable expertise with server-side coding. In order to prevent unnecessary delay with development, a different approach was found. In edX's "Studio", the course authoring-tool, it is possible to insert an IFrame (Inline-frame) into the webpage. An IFrame is an HTML document that is embedded inside another HTML document in a website (Rouse, 2015). This allowed the serious game to be viewable in the course, without having to integrate the servers. In edX studio the IFrame code is already provided, which makes the IFrame integration little more than replacing some attributes in the code. As edX also provides the possibility to add an extra attribute that ensures the (anonymous) user-id of the student is send to the external URL location, individual student progress tracking is also possible without having students to log in twice. The integration is illustrated in Figure 19 and the actual appearance in the MOOC can be seen in Figure 20.

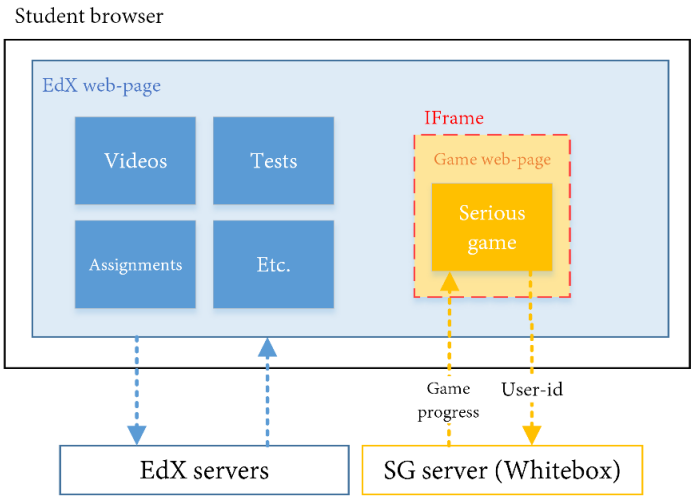


Figure 19: Serious game integration into edX

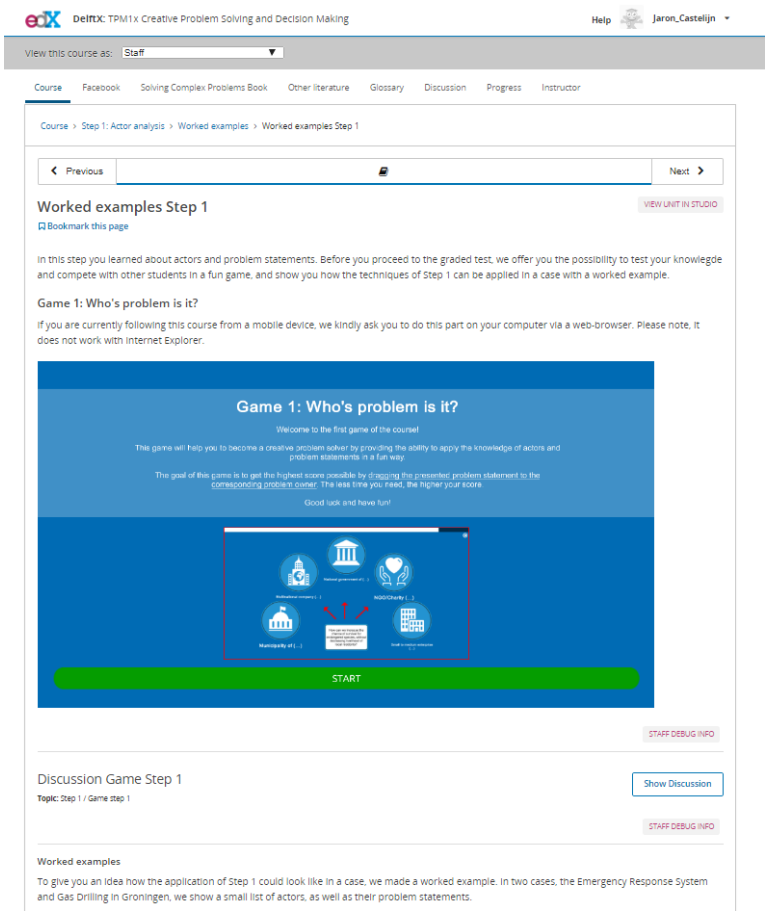


Figure 20: Example of game 1 integration in edX

3.3 Experiment set up

This section describes the experiment set up by discussing the experimental design (3.3.1), the data collection methods (3.3.2), the use sample (3.3.3), the data processing and preparation (3.3.4), and the data analysis methods (3.3.5). To describe the sample and process and prepare the data several descriptive statistics were used. Furthermore, in Sub-section 3.3.4 an Exploratory Factor Analysis (EFA) is conducted to compute factor scores. Computing the factors scores is necessary to be able to use the four latent (e.g. not directly observable) factors to test the hypotheses, and is therefore considered as part of the data preparation and not the main analysis.

3.3.1 Experimental design

To analyse the effects of the serious game in this study, an experiment was used. The experimental design in this study is chosen because of its high internal validity, which enables to determine if the serious game integration was indeed the factor that led or did not lead to the effect on the student retention. There are different types of experimental designs possible, depending on how the experiment can be set-up. The edX platform on which the CSPD MOOC is hosted enables the facilitators to assign students randomly to different cohorts. Furthermore, the content, or treatment, that is accessible by the students can be restricted to one of these cohorts. This enables randomly assigning students to two groups that have access to different content, making it possible to have a *true experimental design*, to be more specific: *a randomized post-test only control group* (Baarda, 2006), visualised in Table 12. The "R" indicates that the students are randomly assigned, the "X" shows which group receives the treatment and represent the dependent variable, and the Y1 and Y2 denote the observations and represent the independent variables.

Table 12: The experimental design for this study

			Post-test
Game group	R	X	Y1
Control group	R		Y2

This experimental design is strong with respect to single-groups threats, multiple group tests, but weak against social interaction threats (Kirk, Kirk, & E., 2003). Single groups occur when there is no comparison of the treatment with another group, and it cannot be said with certainty that the treatment caused the effect on the dependent variable. Multiple group threats occur when the groups are not comparable before the study, and not the treatment but another variable that is different in one of the groups influences the dependent variable. Social interaction threats occur when the two groups are aware that the treatment they get differs. This could lead to diffusion or imitation of treatment (the groups interact and the control group learn from the experimental group), compensatory rivalry (the control group tries to outperform the experimental group), resentful demoralization (control group is demotivated because they don't receive treatment) and compensatory equalization of treatment (facilitators feel sorry for the control group and help them more or provide more resources). However, because MOOCs are followed individually by students from all over the world, and because they can be assigned to different cohorts, the weakness to social interactions threats can be minimized. With the MOOC environment reducing chances of social interaction threats, it can be concluded that the internal validity is high.

3.3.2 Data collection methods

In this study both automated data collection and questionnaires were used as data collection methods. According to Veletsianos and Shepherdson (2016) most papers concerning MOOCs used one (44.8%) or two (38.3%) methods of data collection. Automated data collection was used most (73.2%), followed by questionnaires/questionnaires (55.7%). These two methods are presumably popular as MOOCs provide a rich set of data per student, and questionnaire are found suitable to measure abstract concepts like engagement, motivation and interaction (Hone & El Said, 2016). An overview of the data collection methods, the collected data and for which variable in which part of the theoretical framework the data is used is presented in Figure 21 at the end of this section.

Automated data collection

Automated data collection concerns gathering the trace data, collected by digital platforms. In this study, two digital platforms were present, the edX platform that hosts the CSPD MOOC and the Whitebox platform present at InThere that hosts the game. While the edX platform was used to analyse the engagement and completion of students in the course, the Whitebox was used to analyse engagement with the serious game.

The used dataset provided by edX-platform was the "grade-report", which contained individual (anonymous) data on if the student was in the game group or control group, how many assignments they completed, what grades they obtained for the assignments and if they paid to receive a certificate. This data was used to determine if the student completed the course and see how many assignments they made. Unfortunately it was not possible to have access to the number of videos watched per student and the number of forum posts made per student, as this data resided in too large aggregated datasets that could not be separated in the time span of this research. In case of the forum activity, this was mitigated by counting the forum posts per student manually. Unfortunately, this could not be done for the videos watched per students, meaning the student engagement could only be measured with the number of assignments made and the number of forum posts.

The Whitebox data was used to determine the students' actual serious game engagement. This data was used to support the difference found between the game group and the control group in motivation, interaction, Perceived MOOC Quality, and student retention. As this data was used to support, not all game behaviour data was analysed, only the number of unique games the student played, the total number of games the student played, how many times they played each game, and what their highest score per game was. A students was only considered to have played a game if he or she finished a game, as the game automatically started when the browser was opened and this could not be distinguished from actual gameplay using this dataset.

DelftX questionnaires

As stated in 3.1, the experiment was conducted while the course was already running and it happened that DelftX, who facilitates and coordinates all TU Delft's MOOCs, were already conducting their own research. An implication of this was that the existing pre-questionnaire, mid-questionnaire and post-questionnaire could not be altered or removed. However, some data from these questionnaires could be used to substantiate findings in this research.

The DelftX pre-questionnaire (Appendix 3) contains a combination of 22 qualitative and quantitative closed questions regarding several topics. Interestingly, several questions could be used to collect data on the student level variables: enrolment motivation, intention when enrolling, their proficiency in English, gender, age, and highest obtained degree. Unfortunately, the DelftX pre-questionnaire did not contain data on the prior experience of students or the available hours for the MOOC, of which both was indicated they influence student retention in the theoretical foundation Sub-section 2.2.1. The levels of measurement in the DelftX questionnaire are dichotomous (3), nominal (8), ordinal (8) and ratio (3) depending on the questions and the topic. Despite this questionnaire not specifically being made for this research, it thus contains interesting questions regarding the student demographics to be used as control variables in the theoretical framework parts.

The mid-questionnaire (Appendix 4) only consists of three questions; a five point scale to indicate student's enjoyment halfway through, which could be considered a form of *intrinsic motivation*, and two open questions in which they are asked what their favourite part of the course is so far, and which part needs improvement. These open questions might give insights into how students perceive the game.

The DelftX post-questionnaire (Appendix 5) contained a combination of 44 qualitative and quantitative closed questions regarding the topics: participation level, course perception, instructor quality, forum quality, and content quality. The level of measurement are open (5), nominal (2), ordinal (35), and ratio (2). This data is used for a variety of research by the TU Delft, but also holds interesting questions for this research regarding retention, *extrinsic motivation* (Questions 19 and 21), *interaction* (Questions 5 to 13) and *perceived MOOC quality* (Questions 17, 18, 20, 22, and 25 to 27). In addition, non-completers are asked why they dropped the course. These questions were not specifically made on the basis of this study's theoretical framework, which is why their use is limited. However, they can be used to test the reliability of the findings from the additional questionnaire and substantiate findings where possible.

The additional questionnaires

The additional questionnaire was specifically constructed to measure *motivation, interaction and perceived MOOC quality* and the students' game experience. As the factors are latent, e.g. not directly observable, they were represented with an item-construct relation, see Table 13. The additional questionnaire for the control-group only contained questions related to the four factors, the additional questionnaire for the game-group contained extra questions relating to their game experience. The complete additional questionnaires for both groups can be found in Appendix 6.

Extrinsic motivation was measured by asking the students if they were motivated to complete the course to receive a certificate, to help them grow in their line of work or for course accreditation. Intrinsic motivation is measured by asking if students had fun during the course, found the course interesting, were curious and felt challenged. *Interactivity* is divided into student-student interaction and student-teacher interaction. Unfortunately, no literature was found that provides guidelines to measure either form of interactivity. In this research the interactivity factor is measured by asking students what their feeling of connectedness was with teachers and feelings of

connectedness with other students. *Perceived MOOC quality* will be measured with a questionnaire as well, as it is specifically about the perception of the students. The students will be asked about the main parts that the course consist of: the course materials, assessments, and the overall applicability of the course. An overview of the questions, items and constructs is presented in Table 13.

There exist few systematic approaches to evaluate educational serious games (Petri, Gresse, & Wangenheim, 2016), but those were found too extensive for this research as the number of questions that could be added alongside the DelftX post-questionnaire was limited. Therefore the game experience questions in the additional questionnaire were limited as well, which was not considered as a major limitation as this research is not focussed on the design aspects of the serious game.

Table 13: The factor related questions in the additional questionnaire

Question	**Construct**	**Item**
How much do you agree with the following statement? I remained active in the course because…		
1	*Extrinsic motivation*	I wanted to earn a certificate
2		I wanted to improve my career
3		It relates to my academic program
4	*Intrinsic motivation*	I had fun
5		I found the course interesting
6		The course stimulated my curiosity
7		I felt challenged
8	*Interactivity*	I felt connected with other students
9		I felt connected with the teachers
10		I received enough feedback during the course
11	*Perceived MOOC quality*	There were enough exercises to test my knowledge
12		There was enough course material available
13		The content's applicability to real-world situations is high

Data coupling limitations

During the data collection it appeared that the students' user-ids differed for the automated data collection and the questionnaires. This meant that initially the data from the questionnaires could not be coupled with the grade report, forum activity, and game data. In case of the grade report this could be mitigated by DelftX with equipping the grade report with the same user-ids as the questionnaires. However, this could not be done for the forum activity as this data was collected manually. This had the implication that it was not possible to measure the effect of the four factors on the number of forum posts (theoretical framework part 1), or use the game data to analyse the effect of the serious game on the four factors in more detail (theoretical framework part 2).

Overview of data collection methods and the relation to the theoretical framework parts

In Figure 21 an overview is presented of the data collection methods. The colours indicate which dataset sets are used for the theoretical framework parts. The "supportive" datasets are used to be able to interpret the findings for the theoretical framework in more detail. The videos watched could not be researched as there was no data that could be used. There was useable data for the forum posts but this could not be coupled to the factors from the additional questionnaire.

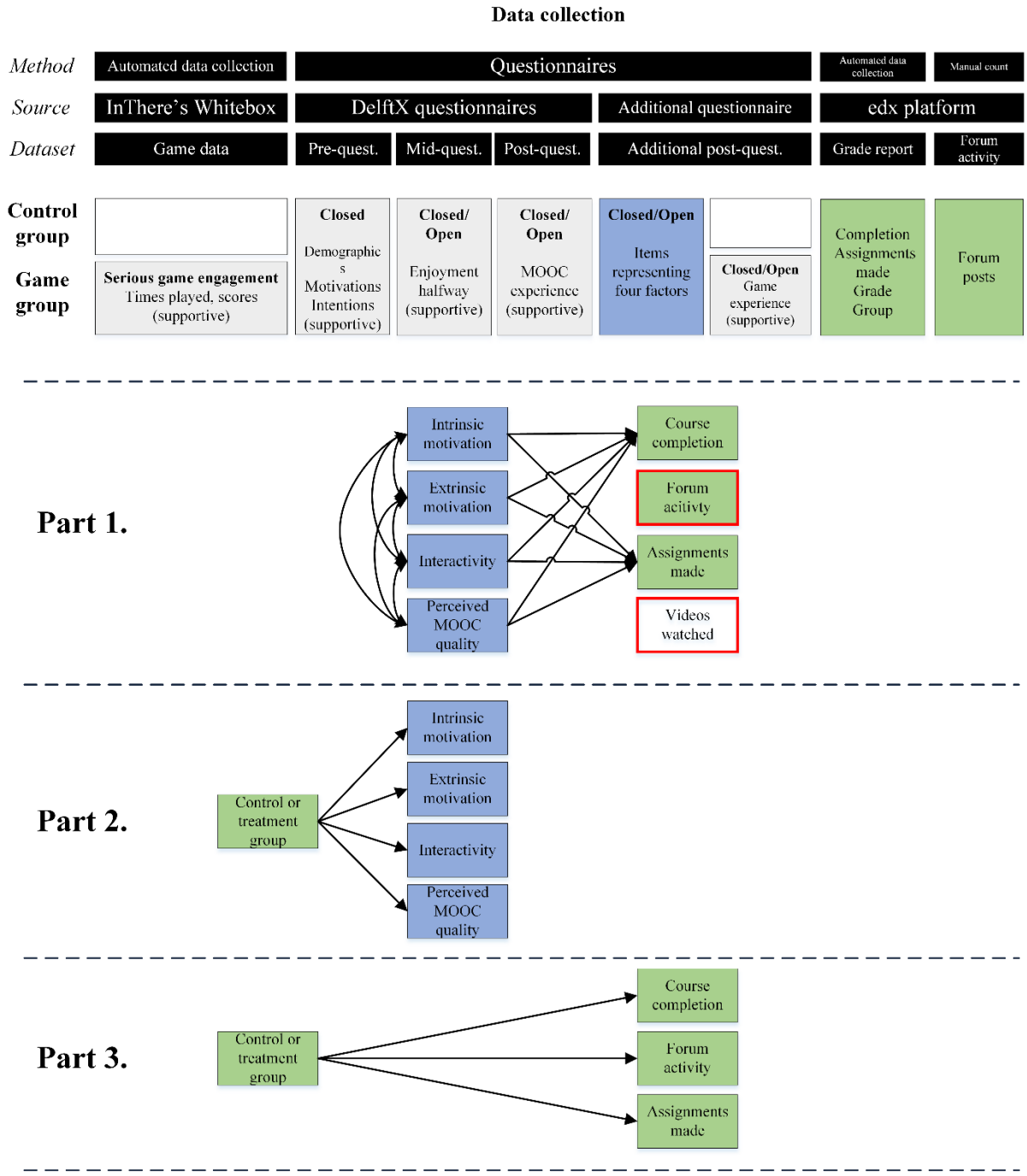


Figure 21: Data collection methods and the relation to the theoretical framework parts

3.3.3 Sample characteristics

The experiment was conducted from the 14th of July to the 5th of September. Initially this research set out to collect data from all students in this period (N = 6135) via automated data collection, manual counting and questionnaires. While with the automated data collection and manual counting data of every student in both groups (the MOOC facilitator user-IDs were taken out) was obtained, the questionnaires suffered a high non-response (Table 14).

Table 14: Questionnaire responses and percentage of responses by completers

Questionnaires	All students	Started students	Of which completers	Completers / all students (%)	Completers / started students (%)
Grade report, Game data, forum report	6135	254	73	1.2%	28.7%
DelftX pre-questionnaire	362	97	29	8.0%	30.0%
DelftX mid-questionnaire	46	45	30	65.2%	66.7%
DelftX post-questionnaire	33	31	20	60.6%	64.5%
Additional questionnaire	58	52	39	67.2%	75.0%

Counting students as started

When analysing the responses further, it was discovered that 96% out of the 6135 students did not finish any assignment. This does not mean that these students did not interact with the course at all, of these students some engaged with the serious game (n = 17) or filled in DelftX's pre-questionnaire (N = 271), but indicates that a large group of students lacked the motivation to complete one step. This high number influences the completion rate severely. DelftX accounts for this in their evaluation reports by also showing the completion rates of students who at least clicked on one item in the courseware. However, this is not deemed an appropriate way of calculating the completion rate, as students who are just checking out the course presumable click on items in the courseware and are therewith included in the completion rate. It was also not possible to restrict the sample size to that of students who indicated in the pre-questionnaire they intended to complete the course, as not all these students filled in the other questionnaires and the already small sample size would then become even smaller. Instead, in this study it was decided to count students as *started* students when they finished at least one assignment in the course, as this makes it more likely the students intend to complete the course, making the completion rate a more valuable statistic. This is also in accordance with authors proposing to determine the completion rate excluding students who did not intend to complete the course (Koller et al., 2013; Wilkowski et al., 2014). Lastly, due to the absence of track data indicating when students enrolled for the course, the completion rate was lowered by students who might have signed-up 1 day prior to the data extraction for this research. This means that students who did not have the time to finish the course, were counted as non-completers. This limitation made it difficult to compare the completion rate related findings in this MOOC to other MOOCs.

Self-selection bias

It is important to note that the sample cannot be considered random in case of the mid- post and additional questionnaire, as the number of completers that filled in these questionnaires was found to be significantly higher than the number of non-completers. This means there is a self-selection bias that has a large impact on the extent to which the questionnaire data can be generalized to the MOOC population.

Sample characteristics

Only the started students were included in the sample. As the questions regarding the student demographics resided in the DelftX pre-questionnaire, the student demographics are described based on the 97 responses and are assumed to be representative for the 254 students in total in this sample. The average age of the students was 36 (SD = 13.6). When analysing the gender it was found that 57 (63%) students were male and 33 (37%) were female. 18 (21%) students had a professional degree or lower, 33 (38%) had a bachelor degree, and 35 (41%) a master's degree or higher.

12 (14%) participants indicated their current occupation was student, 1 (1%) recently graduated, 62 (69%) were working professionals, 7 (8%) were looking for a job, and 1 (1%) was a parent or care giver,
4 (5%) were retired and 2 (2%) indicated another occupation.

Regarding the continent of origin it was found that 32 (36%) students were from Europe, 25 from Asia (28%), 14 (15%) from North America, 9 (10%) from South America, 8 (9%) from Africa, and 2 (2%) from Oceania/Australia.

When analysing the enrolment motivations it was found that 30 (73%) students indicated they enrolled to use the course in their daily work activities or to improve their future career prospects and 25 (27%) students enrolled because they were interested in this subject and wanted to learn more about it,

Lastly, 80 (87%) of the students indicated that they intended to complete the course, while 12 (13%) did not intend to complete the course. This means that 12 students indicated they did not intended to complete the course in the pre-questionnaire, but did end up making the first assignment and were considered started students in this study.

3.3.4 Data processing and preparation

This section explains how the collected data was processed and prepared in order to find the results in Chapter 4. This started with transforming the variables from the different datasets into numeric variables following a coding scheme, so that the data could be analysed with the SPSS statistical software package. The datasets were then checked on mistakes or errors with descriptive statistics. Lastly, a factor analysis was used to compute the student scores on the four latent factors, based the item-construct operationalization in the questionnaires. The results of the descriptive statistics and factor analysis are shown in this chapter, as they are not results used to answer the research questions in Chapter 4, but are results used to prepare the data for the main analyses regarding the theoretical frameworks.

Data coding scheme

The data from all six datasets had to be recoded into workable numbers before they could be analysed. To keep a clear overview of the data, a prefix was added before every variable code that indicated the source. There are *tpeq.*, *tmq.*, *tpoq.*, *aq*. for the DelftX pre-questionnaire, DelftX mid-questionnaire, DelftX post-questionnaire and additional questionnaire respectively. The prefix *gr.*, *ga,* and *fo.* were used to indicate the grade report game dataset and forum data. Furthermore, the specific variables that are used as items for the *Motivation*, *Interaction* and

Perceived MOOC Quality factors were coded with an additional prefix that indicates what construct they measure. The items that were used to measure extrinsic motivation have the prefix *em*, intrinsic motivation with *im*, interaction with *int*, and perceived MOOC quality items with *pmq*. The complete data coding scheme can be found in Appendix 7.

Data quality

Before the collected data was analysed, a frequency analysis was conducted to check for errors. The frequency analyses showed that two students in the game group were only present in edX's grade report, and not in InThere's game dataset due to the different import date of both files. This was corrected by updating the grade report to one with a matching import date to the edX dataset. Furthermore, it was found that the student who responded with a five on the 5-point-likert scale questions in the additional questionnaire, were not present in the dataset at first, which was corrected.

The frequency analysis also showed that for the questions regarding the serious game design missing values were seen as zeros. It was also found in the question asking about students' enrolment motivations (N = 360), two categories were only mentioned once by a student: "I want to get a degree in this field" and "I know the instructors and am interested in their work". These categories were removed from the dataset. For the student level variable "Highest obtained degree", it was found not all categories had enough respondents. It was therefore decided to recode these variables into three categories: professional degree or lower, bachelor degree, and master degree or higher. Also the enrolment intention of students was recoded into two variables: "I intend to complete" and "I do not intent to complete", as there were too few responses for the other categories. Despite the Current occupation variable and Continent variable having low responses for some categories as well these categories were not recoded as it was not possible to group these categories (e.g. you cannot group students from Oceania and Africa, or students occupations like "parent / care giver" with "retired").

To conclude the data quality assessment the factor scores were visualised in a boxplot to identify outliers. This was done as outliers could indicate coding mistakes or means one of the observations was fundamentally different from the others. The boxplot and interpretation of the outliers are shown in Appendix 8. It was concluded that there was no reason to exclude the outliers from the dataset.

Factor analysis

Before the data was analysed, the students' scores for each of the four factors had to be computed. It was chosen to compute factors scores for each student with a factor analysis. According to Matsunaga (2010) an Exploratory Factor Analysis (EFA) should be used when researchers are unsure of how variables would operate with one another, and a Confirmatory Factor Analysis (CFA) should be used when researcher have a clear theory about the underlying structure. As stated in Sub-section 2.3.2, there is an theoretical basis for each factor individually, but not for the relation between the variables. Therefore and EFA was deemed most appropriate for this study to validate whether the operationalization of the factors with the construct-item relation is done appropriately, and to compute the factor scores for each student. Matsunaga states that in order to perform an EFA the researches should strive for as large a dataset as possible, with an advised minimum of n = 100. In this study the sample was n = 58

(additional questionnaire response), which is significantly lower than the advised minimum, meaning the results of the factor analysis should be interpreted carefully.

Note that in this study the 5-point Likert scale items are considered of interval level, not ordinal. Whether or not this is possible is an ongoing debate in scientific literature. The decision in this paper is based on the journal paper of Norman (2010), that has a 24.12 field weighted citation impact on Scopus. He argues that the use of parametric methods is often unfoundedly challenged due to the ordinal nature of Likert-scales. He states that studies dating back to the 1930 has shown the high robustness of central tendency tests like the ANOVA, regression and correlation methods like the Pearson correlation. This high robustness means that even if the assumptions are violated, in this case that the measurement levels need to be ordinal, the tests will still provide the right answer and can be utilized without concern for getting the wrong answers (Norman, 2010).

The factor analysis in this study was performed in accordance with Matsunaga's (2010) recommendations. The factor extraction method was Principal Axis Factoring, the rotation method was oblique-rotation. He advises the use of an oblique-rotation method as the orthogonal rotation method suggest that the factors are unrelated, which is rarely the case in reality in social sciences and is also not considered likely in this theoretical framework as these relations are specifically researched. Instead of determining the number of factors based on an eigenvalue greater than one, a Parallel Analysis was used to force 4 factor solutions in accordance with the theoretical framework and Matsunaga's recommendations.

The factor model is better the more a "simple structure" is represented. The simple structure is represented more accurately when each variable correlates (loads) highly on one identified factor, and low on the others. The factor loading is the direct causal effect of the factor on the indicator and can be found in the *Pattern Matrix*. When an oblique rotation method is used, as in this case with the oblimin rotation, the factor loadings in the pattern matrix are regression coefficients. What is considered a high enough and or too low factor loading is debated amongst authors, and a factor analysis often requires a high amount of interpretation from researchers (Matsunaga, 2010). Before the final result presented in Table 15 was obtained, it was found that the interactivity variable "I felt connected with other students" distorted the factor analysis and prevented finding a solution, which is why it was removed. After that, it was found that the perceived MOOC quality variable "The content's applicability to real-life situations is high", loaded about the same on the first (0.460) and fourth (-0.471) variable and was also removed to obtain a simple structure. In the final Pattern Matrix in Table 15 it can be seen that after removing the applicability indicator, only the "I had fun" indicator loads highly on two factors. Since there is quite a difference between the loadings and there is a strong theoretical basis for "fun" being part of intrinsic motivation, it was decided not to exclude this variable from the factor analysis.

Table 15: Pattern Matrix with Crobach's Alpha

	Perceived MOOC quality	**Interactivity**	**Extrinsic motivation**	**Intrinsic motivation**
Crobach's Alpha	*0.787*	*0.791*	*0.695*	*0.855*
There were enough exercises to test my knowledge	1.014			
There was enough course material available	0.741			
I felt challenged	0.487			
I felt connected with the teachers		1.013		
I received enough feedback during the course		0.586		
I wanted to improve my career			0.723	
I wanted to earn a certificate			0.618	
It related to my academic program			0.531	
The course stimulated my curiosity				-0.947
I found the course interesting				-0.697
I had fun		0.428		-.636

It can be concluded that the pattern matrix shows a simple structure, indicating that the indicator variables derived from the questionnaire indeed measure the presumed latent factors perceived MOOC quality (factor 1), interactivity (factor 2), extrinsic motivation (factor 3), and intrinsic motivation (factor 4). Interestingly enough it can be seen that the item originally intended to measure intrinsic motivation, challenge, is significantly more influenced by perceived MOOC quality than intrinsic motivation. That means students who answered that they continued the course because the felt challenged, saw this as an important quality of a MOOC. As this is not an unthinkable line of reasoning, it was decided to keep this variable as item for perceived MOOC quality instead of for intrinsic motivation.

Also, it can be seen that the intrinsic motivation variables have a negative factor loading. It is debated why the factor scores are presented with negative correlations, but it is agreed upon that the factor loadings and scores are indeterminate, meaning that the factor loadings can interpreted as absolute values without their negative or positive connotation (Steiger & Schönemann, 1978). A consequence of SPSS "choosing" a solution with a negative factor loading, is that when the factors analysis is used to compute a factor score, the factor score is reversed. This means that when students have a high factor score, their intrinsic motivation is low and vice versa. Note that it does not mean that the students who score highly on intrinsic motivation score low on the other factors, as the correlation amongst the factors is also negative (see factor correlation matrix in Appendix 9).

Lastly, it can be seen in the pattern matrix that the items "There were enough exercises" and "I felt connected with the teacher" had a factor loading greater than 1. This is possible as the factor loadings in the Pattern Matrix represent regression coefficients and not correlation coefficients, but do indicate that the factor is highly influenced by this variable and vice versa.

With an obtained simple structured and acceptable high Crobach's alpha, it was concluded that the reliability of the factors operationalization in the post-questionnaire was acceptable and the factor scores for extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality could be computed and used in the further analysis. Note that the variables are standardized, meaning that if a student has a score close to 0, the students score closely to the average of the sample.

3.3.5 Data analysis methods

This section describes which data analyses methods were used to obtain the results in Chapter 4 and why this method was deemed appropriate. This section is structured following the three parts of the theoretical framework that will be evaluated. For each part, there is a distinction made for data analysis methods used for the *main statistical analysis* and the *supportive statistical analysis*. The main statistical analysis concerns analyses that are used to test the hypotheses and directly evaluate the framework, while the supportive statistical analysis concerns data that was gathered to provide context for the framework evaluation.

Data analysis methods for evaluation of theoretical framework part 1

The analyses that were conducted to evaluate theoretical framework part 1 are the relations between the factors and the course completion, the relation between the factors and the number of assignments made, and the relations amongst the variables themselves. The supportive analysis for the evaluation of the theoretical framework part 1 concerned the relation between course completion and the answers to the questions in the DelftX mid-questionnaire and DelftX post-questionnaire that somewhat resembled the questions used to measure the four factors in the additional questionnaire.

Factors vs. course completion: independent sample t-test & Whitney-Mann U-test

The factors are of interval measurement level and course completion is of nominal measurement level. The difference in factors scores between students who completed the MOOC and students who did not should be analysed with an independent sample t-test. However, A Shapiro-Wilk's test and a visual inspection of their histograms, normal Q-Q plots and box plots (Appendix 10) showed that the factor scores were not approximately normally distributed for both completers and non-completers. Whether or not a violation of the basic assumption of normality should be a reason to use a nonparametric test instead of a parametric test like the t-test is an ongoing debate in scientific literature. Norman (2010, p. 628) states that it has been shown with both data and theory that "parametric methods examining differences between means, for sample sizes greater than 5, do not require the assumption of normality, and will yield nearly correct answers even for manifestly non-normal and asymmetric distributions like exponentials".

Originally it was intended to evaluate the theoretical framework part 1 using the control group only, to prevent distortion from half of the students having access to the serious game. However, it was found that only four students in the control group completed the course and filled in the additional questionnaire, meaning that the non-completer group would consist out of four students. As this is lower than the threshold of N > 5 stated by Norman (2010), this meant a nonparametric Whitney-Mann U-test should be used. If the use of the smaller more homogeneous sample with a nonparametric test would be better than a larger more heterogeneous sample was debatable, which is why it was decided to use both the independent t-test and the Mann-Whitney U-test and compare the results.

Factors vs factors: Pearson correlation

The linearity for the correlations between each factor were visually assessed with scatterplots. The scatterplots for the factors scores can be found in Appendix 11. It was difficult to distinguish any correlations, be it linear or non-linear, so there was no reason found to refrain from using the Pearson correlation test.

Factors vs assignments made: not possible

It was not possible to compute the Pearson correlation between the four factors and the number of assignments made. This was due the factor scores being obtained from the additional questionnaire, which was mostly filled in by completers of the MOOC (see Sub-section 3.3.3), and almost all of the students who completed the MOOC made all 5 assignments. This meant that of the students whose factor scores could be compared to the number of assignments they made, 93% made all assignments.

DelftX factor questions vs completion: independent sample t-tests (supportive)

Several questions in the DelftX questionnaires somewhat resembled the items used to measure the four factor constructs. The answers possibilities followed a 5-point Likert scale, and they were therefore regarded as interval measurement level. This meant an independent sample t-test should be used to compare the means of these questions between completers and non-completers.

Data analysis methods for evaluation of theoretical framework part 2

The analyses that were conducted to evaluate theoretical framework part 2 concerned the difference between the serious game group and the control group concerning the four factors. There were several supportive analysis used to provide context for this evaluation. First it was analysed if there was a difference between the serious game group and the control group in the answers to the DelftX factor related questions. Second, it was analysed if the random allocation by edX led to two identical groups by comparing the student demographics of both groups. Third, the quantitative and qualitative data from the serious game evaluation questions and forum discussions were analysed.

Serious game vs factors: independent sample t-test

As two groups are compared, the serious game variable is of nominal measurement level. The factor scores are of interval measurement level. This means the influence of the serious game should be measured by comparing the means with an independent sample t-test. However, a Shapiro-Wilk's test and a visual inspection of their histograms, normal Q-Q plots and box plots (Appendix 12) showed that the factor scores regarding perceived

MOOC quality, interactivity, and extrinsic motivation were also not approximately normally distributed for either the control group or game group. As explained earlier, it has been shown that the t-test is robust against violations of the normal population assumption (Norman, 2010) for samples with N > 5. Since this was the case for both the game (N = 22) and control (N = 30) group, it was decided that an independent sample t-test could still be used.

Serious game vs DelftX factor questions: independent sample t-test (supportive)
As stated in the data analysis methods description for theoretical framework 1, the DelftX factor questions are of interval measurement level, which is why an independent sample t-test was used.

Serious game vs student demographics: independent sample t-test, contingency tables with chi-square tests, and ANOVA (supportive)
The student demographics that were compared between the groups were: if the students payed for a certificate or not, age, English proficiency, gender, highest obtained degree, current occupation, continent of origin, enrolment intention and enrolment motivation. These were of nominal, interval, and ratio levels, meaning independent sample t-tests, contingency tables with chi-square, and ANOVA tests were used.

Quantitative serious game evaluation: descriptive statistics (supportive)
The quantitative serious game evaluation contained questions that were measured with a 5-point Likert scale. As these questions were only present for the game group, descriptive statistics were used to assess what students thought of the game.

Qualitative serious game evaluation: no statistical method (supportive)
The qualitative serious game evaluation responses stemmed from open questions in all questionnaires and discussions of the game on the discussion forum. This responses were grouped and interpreted without further statistical analysis.

Data analysis methods for evaluation of theoretical framework part 3
The data analysis methods for the third part of the theoretical framework evaluation concerned the difference between the game group and control group regarding the number of completers, number of assignments made, and number of forum posts made. Two supportive analyses were conducted in addition to the supportive analyses from the theoretical framework part 2. These were the differences between the groups regarding the grade they obtained, and the number of students in the game group who played a game in each step. The last analysis only required descriptive statistics and is not further elaborated upon in this section.

Serious game vs completion: contingency table with chi-square
Both variables are of nominal measurement level, meaning a contingency table should be computed and a chi-square test had to be conducted to test for the significance of the findings.

Serious game vs number of assignments, forum posts, and grade: independent sample t-tests

The serious game variable is of nominal level, and the number of assignments, forum posts, and grade are of ratio level. This meant independent sample t-tests should be used. No extensive normality checks were performed, as a simple visual check of the distributions (Appendix 13) clearly showed there was no normal distribution. However, the sample sizes were large enough to comply with Norman's (2010) threshold of $N > 5$.

3.4 Revised hypotheses

Due to several limitations during the data collection, not all hypotheses that this research aimed to test could be answered. The effect of the four factors on course engagement could not be researched, as the forum activity could not be coupled to the questionnaire dataset, no data could be used on the number of videos watched per student, and 93% of the students of whom their factor scores were obtained made all 5 assignments. This means that for the theoretical framework part 1, student retention was analysed solely on the completion rate, and not included the student engagement. Furthermore, the hypotheses for part 2 and 3 are formulated differently, now that the experimental design made clear a difference between the groups is analysed.

Revised hypotheses for evaluation of theoretical framework part 1

H1.1	Students who completed the course have a higher intrinsic motivation than students who did not complete the course.
H1.2	Students who completed the course have a higher extrinsic motivation than students who did not complete the course.
H1.3	Students who completed the course have experienced more interaction than students who did not complete the course.
H1.4	Students who completed the course have a higher perceived MOOC quality than students who did not complete the course.

Revised hypotheses for evaluation of theoretical framework part 2

H2.1	Students in the game group have a higher intrinsic motivation than students in the control group
H2.2	Students in the game group have a different extrinsic motivation than students in the control group
H2.3	Students in the game group have a higher interactivity than students in the control group
H2.4	Students in the game group have a different perceived MOOC quality than students in the control group

Revised hypotheses for evaluation of theoretical framework part 3

H3.1	More students in the game group complete the course than students in the control group
H3.2	Students in the serious game group make more assignments than students in the control group
H3.3	Students in the serious game group make more forum posts than students in the control group

3.5 Chapter summary and conclusions

This chapter described the experiment that was used to test the hypotheses from Chapter 2. The experiment was conducted in the Creative Problem Solving and Decision Making MOOC facilitated by DelftX on the edX platform. This MOOC teaches an analytical approach to solve complex problems in five steps. In each of these steps several analytical methods are used and related to the previous and following steps. These findings were taken into account with the design, development and integration of five serious games. The serious games were closely related to the methods learned in the CSPD MOOC and utilized game mechanics and textual encouragement to influence the intrinsic motivation, extrinsic motivation, interaction, and perceived MOOC quality of students. The serious games were incorporated into the MOOC via an IFrame.

Several questionnaires and automated data collection methods were used to collect data for the theoretical framework evaluation. All questionnaires were filled in by significantly more students who completed the MOOC than students who did not complete the MOOC, indicating a selection bias. Furthermore, 96% of the students did not finish at least one assignment, which led to the decision of excluding these students when analysing the student retention in the MOOC, reducing the sample size from 6135 to 254. The intrinsic motivation, extrinsic motivation, interactivity and perceived MOOC quality factors were measured with an item-construct relation, and the validity of this operationalisation was checked with a factor analysis. The factor analysis led to the removal of two items before it could be concluded that the item-construct operationalisation indeed measured the four factors, and the factors score for every student for every factor could be computed an used to evaluate the theoretical framework part 1 and 2. Lastly, revised hypotheses were formulated as the students' course engagement could not be analysed when evaluating theoretical framework part 1. With the new hypotheses and all data prepared, the results in Chapter 4 could be obtained.

4. Results

In this chapter the results obtained from the experiment are analysed and used to answer research questions 3, 4 and 5. Each of these research questions are answered by testing the hypotheses of theoretical framework part 1, part 2, and part 3.

3. *To what extent do extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality influence student retention?*
4. *To what extent does a serious game affect extrinsic motivation, intrinsic motivation, interactivity, and perceived MOOC quality?*
5. *Is there a difference in student retention when a serious game is incorporated into a MOOC?*

First, the effects of the factor on student retention is assessed in Section 4.1, before the effect of the serious game on the factors is assessed in Section 4.2, and the effect of the serious game on the student retention is assessed in Section 4.3. Including the supportive analyses mentioned in Section 3.3.5, the detailed chapter structure is as follows:

4.1 Part 1: The factor effects on student retention

The effect of the four factors on student retention was assessed by comparing the means for the factors between students who completed the course and students who did not. The correlation amongst the factors was analysed as well.

4.1.1 Main statistical analyses results part 1

Factors vs course completion

No significant differences were found for any of the factors scores when students who completed the course were compared with students who did not, using an independent sample t-test (completed N = 39, not completed *N* = 13), see Table 16, or using a Whitney-Mann U-test (completed N = 26, not completed *N* = 4), see Table 17.

Table 16: Factor score comparison between completers and non-completers using t-test

Factor	Variance assumption	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
IM	Equal variances assumed	0.874	0.354	0.605	50	0.548	0.181	0.300
EM	Equal variances not assumed	6.018	0.018	-0.419	14.927	0.681	-0.155	0.371
INT	Equal variances assumed	0.683	0.412	0.27	50	0.789	0.086	0.318
PMQ	Equal variances not assumed	8.749	0.005	-0.631	14.836	0.537	-0.246	0.389

Table 17: Factor score comparison between completers and non-completers using Mann-Whitney U-test

	IM	EM	INT	PMQ
Mann-Whitney U	33.000	37.000	40.000	32.000
Wilcoxon W	43.000	47.000	391.000	383.000
Z	-1.159	-0.915	-0.732	-1.220
Asymp. Sig. (2-tailed)	0.246	0.360	0.464	0.222

Factors vs factors

The correlations between the factors were computed only for the control group (N = 30) and are shown in Table 18. The only moderate positively significant correlation (r = 0.402) was found between extrinsic motivation and interactivity. Note that a high intrinsic motivation factor score indicates a low intrinsic motivation due to the factor score being reversed when computed with the factor analysis. This means the (non-significant) relations between intrinsic motivation and the other three factors can be regarded as positive.

Table 18: Pearson correlation between factors

		IM	EM	INT	PMQ
IM	Pearson Correlation	1	-0.269	-0.230	-0.094
	Sig. (2-tailed)		0.151	0.221	0.620
	N	30	30	30	30
EM	Pearson Correlation	-0.269	1	0.402*	0.024
	Sig. (2-tailed)	0.151		0.028	0.902
	N	30	30	30	30
INT	Pearson Correlation	-0.230	0.402*	1	0.052
	Sig. (2-tailed)	0.221	0.028		0.785
	N	30	30	30	30
PMQ	Pearson Correlation	-0.094	0.024	0.052	1
	Sig. (2-tailed)	0.620	0.902	0.785	
	N	30	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

4.1.2 Supportive statistical analyses part 1

DelftX questions resembling the four factors vs completion

Several questions in the DelftX questionnaires somewhat resembled the items used to measure the four factor constructs. When these questions were analysed, again no significant differences were found for any of the questions between completers and non-completers. The questions can be found in Appendix 4 and 5, the statistical results can be found in Appendix 14.

4.1.3 Summary of findings theoretical framework evaluation part 1

This section analysed if intrinsic motivation, extrinsic motivation, interactivity and perceived MOOC quality did indeed influence student completion, as shown in the literature study in theoretical foundation chapter. The influence of the factors was analysed in a sample containing only the control group with a Mann-Whitney U test and in the sample containing both the game and control group with an independent sample t-test. With both tests no significant differences were found when comparing the factors scores between students who completed the MOOC and students who did not complete the MOOC.

When analysing the questions in the DelftX that resembled the question items used to measure the four factors in the additional questionnaire, no differences were found either. Therefore it can be concluded that there is no relation between any of the factors and course completion, meaning none of the hypotheses for the theoretical framework part 1 are supported, see Table 19.

Table 19: Result revised hypotheses theoretical framework part 1

H1.1	Students who completed the course have a higher intrinsic motivation than students who did not complete the course.	Not supported
H1.2	Students who completed the course have a higher extrinsic motivation than students who did not complete the course.	Not supported
H1.3	Students who completed the course have experienced more interaction than students who did not complete the course.	Not supported
H1.4	Students who completed the course have a higher perceived MOOC quality than students who did not complete the course.	Not supported

When analysing the correlations between the factors, there was only a moderately positively significant correlation found between extrinsic motivation and interaction. This means that the more the students felt they interacted with the teachers (student-student interaction was removed in the factor analysis, see Sub-section 3.3.4), were more likely to have extrinsic motivations for completing the course.

4.2 Part 2: The serious game effect on the factors

The effect of the serious game on the factors was assessed by comparing the means for every factor between the students in the control group ($N = 30$) and students in the game group ($N = 22$). The sample size was restricted to students of whom the factor scores could be calculated.

4.2.1 Main statistical analysis part 2

Serious game incorporation vs four factors

Students in the serious game group reported a *higher* intrinsic motivation *factor score* than students in the control group. With the factor score for intrinsic motivation being reversed, this meant that students in the game group had a *lower* intrinsic motivation (see *factor analysis* in Sub-section 3.3.4). No significant differences were found when comparing the means of extrinsic motivation, interactivity and perceived MOOC quality between the two groups, see Table 20 and Table 21.

Table 20: Mean factor score comparison between groups

	Group	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
IM	Control group	30	-0.25	0.76	0.14
	Game group	22	0.28	1.06	0.23
EM	Control group	30	-0.00	0.75	0.14
	Game group	22	0.02	1.08	0.23
INT	Control group	30	-0.12	0.94	0.17
	Game group	22	0.16	1.04	0.22
PMQ	Control group	30	0.13	0.73	0.13
	Game group	22	-0.19	1.16	0.25

Table 21: Independent sample t-test for factors scores and game and control group

Factor	Variance assumption	**F**	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**
IM	Equal variances assumed	3.441	0.069	-2.102	50	0.041	-0.53147	0.25279
EM	Equal variances assumed	0.836	0.365	-0.118	50	0.906	-0.03003	0.254044
INT	Equal variances assumed	0.007	0.932	-1.003	50	0.321	-0.27676	0.275869
PMQ	Equal variances not assumed	6.190	0.016	1.174	32.919	0.249	0.328818	0.280002

4.2.2 Supportive statistical analysis part 2

Serious game incorporation vs DelftX questions resembling the four factors

The DelftX questions that were used to support the findings for the theoretical framework part 1 evaluation were also used to support the analysis of part 2. The question somewhat resembling the intrinsic motivation stemmed from the DelftX pre-questionnaire, in which students were asked to rate their enjoyment halfway through the course on a scale from 1 to 5. However, no significant differences between the control group and game group were found for this question or any of the questions resembling the other three factors. All the others questions and the statistical results can be found in Appendix 14.

Comparing groups with student demographics

The two groups were compared in terms of student demographics to see if there were any other differences between the groups than the serious game intervention. The complete comparison can be found in Appendix 15. A significant difference was found for the age ($t = -2.39$; $df = 87$; $p < 0.019$), showing that the students in the game group were on average 6.8 (SE = 2.8) years older. No significant differences were found between the groups for the number of paying students ($p < 0.525$), their English proficiency ($p < 0{,}095$), gender ($p < 0.659$), highest obtained degree ($p < 0.774$), enrolment intention ($p < 0.756$), and enrolment motivation ($p < 0.334$). The groups could not be compared regarding the current occupation and continent of origin of the students, as a number of the cells in the contingency tables had an expected count less than 5, which violates the assumption for a chi-square test (Agresti & Finlay, 2009).

When tested, age was not found to have a significant correlation with any of the four factor scores, as shown in Table 22. This quantitative analysis could be substantiated to some extent with findings from the DelftX post-questionnaire, which contained several questions that could be relevant regarding the game influence, as the games could be seen as “quizzes”. However, no significant differences were found between the serious game group and control group, as shown in Appendix 16**Error! Reference source not found.** (control group N = 12, game group N = 8).

Table 22: Correlation between age and factors

		IM	EM	INT	PMQ
Age	Pearson Correlation	-0.021	-0.199	-0.058	-0.026
	Sig. (2-tailed)	0.926	0.375	0.798	0.909
	N	22	22	22	22

Quantitative game evaluation

The serious game was evaluated in the additional questionnaire with specific questions, of which the answers are shown in Table 23. All questions appeared to be answered positively, including the questions regarding *fun* and *curiosity*, which are two items used for the intrinsic motivation factor construct that was found to be lower in the game group.

Table 23: Descriptive statistics of answers to quantitative game evaluation questions

Question	N	Min.	Max.	Mean	Std. Deviation
The games motivated me to continue with the course	20	1	5	3.50	0.889
The games were useful	19	1	5	3.58	1.071
The games were fun	19	1	5	3.37	1.012
The games were challenging	19	1	5	3.63	0.955
The games reminded me why I enrolled for this course	19	1	4	3.32	0.885
I was curious what the next game would be	20	1	4	3.20	1.005
I tried to get a higher rank in the games	20	1	5	3.65	0.988
I studied extra because I wanted a better multiplier	19	1	5	3.37	1.012
The games helped me to understand the course material	20	1	5	3.55	1.234

Qualitative game evaluation

There were also several ways qualitative feedback about the game was collected: via the open questions in the DelftX mid-questionnaire, DelftX post-questionnaire, the additional questionnaire and via the discussion forum. The amount of qualitative data was limited, which was expected given the quantitative focus of this research, and is shown in Appendix 17. It could be derived from these responses that the game was mostly regarded as positive, but the timing of the games needed to be improved as this was a much mentioned negative factor.

4.2.3 Summary of findings theoretical framework evaluation part 2

This section showed if there was a difference in intrinsic motivation, extrinsic motivation, interactivity and perceived MOOC quality between students in the game group and students in the control group. An independent sample t-test showed that the students in the game group had a lower intrinsic motivation than students in the control group.

When looking at the students' course enjoyment halfway through the course, no significant difference was found between the two groups. However, as it is debatable whether enjoyment can be considered strictly intrinsic and this measurement was taken halfway through the course, it was not seem as enough reason to dispute the lower intrinsic motivation result in the game group.

Interestingly, it was found that despite the random allocation of students by edX to either group, the students in the game group were on average 7 years older. However, age was not found to have a significant correlation with intrinsic motivation or any other factor.

The quantitative game evaluation showed that the students regarded the game as moderately positive for all questions, including the questions related to intrinsic motivation. The qualitative game evaluation showed that the game was both perceived as positive and negative by students, and that timing of the games seemed responsible for most negative responses.

The main analysis and supportive analysis led to the conclusion that the lower intrinsic motivation could be attributed to the presence of the game, with a strong suspicion that it was caused by students getting frustrated with the lack of time provided for a question in the each game based on the qualitative questionnaire. With the lower intrinsic motivation for the game group, the null hypothesis for intrinsic motivation was rejected, but not because the intrinsic motivation was higher, as was expected, but because it was lower. The alternative hypotheses for extrinsic motivation, interactivity, and perceived MOOC quality were not supported either, as no significant differences were found between the game group and control group, see Table 24.

Table 24: Results revised hypotheses theoretical framework part 2

H2.1	Students in the game group have a higher intrinsic motivation than students in the control group	Not supported, was lower
H2.2	Students in the game group have a different extrinsic motivation than students in the control group	Not supported
H2.3	Students in the game group have a higher interactivity than students in the control group	Not supported
H2.4	Students in the game group have a different perceived MOOC quality than students in the control group	Not supported

4.3 Part 3: Serious game effect on student retention

The effect of the serious game on the student retention was assessed by analysing the difference between the control group and game group in the number of students who completed the course, the number of assignments made and the number of forums posts by students. The data for this research stemmed from the grade report and manual forum post count, meaning that the sample consisted of all started students ($N = 254$).

4.3.1 Main statistical analyses results part 3

SG vs completion

No significant difference ($X^2 = 0.110$; df =1; $p < 0.740$) was found between the number of completers in the control group and in the game group, see Table 25.

Table 25: Number of completers in game group and control group

			Completed	**Not completed**	**Total**
Group	Control group	Count	42	100	142
		% within Group	29.6%	70.4%	100.0%
	Game group	Count	31	81	112
		% within Group	27.7%	72.3%	100.0%
Total		Count	73	181	254
		% within Group	28.7%	71.3%	100.0%

SG vs assignments made

No significant differences ($F = 0.346$; $t = -1.01$; $df = 252$; $p < 0.424$) were found between the control group and the game group in the total number of assignments made per student. When analysing the number of students who made each assignment in each group, no significant differences were found either, see Table 26. The fact that assignment five was made by more students in both groups than assignment four could be explained by the fact that the course was self-paced and students did not have to follow a specific order of assignments.

Table 26: Number of started students who made each assignment, per group

Assignments	Total	Control group	Game group	p <
Assignment 1	243	138	105	0.249
Assignment 2	145	75	70	0.306
Assignment 3	118	64	54	0.704
Assignment 4	102	53	49	0.128
Assignment 5	105	54	51	0.222

In Table 26 it can be seen that about the same number of students in both groups made all assignments, while less student started in the game group. Approximately 46% of the students who started in the game group made all 5 assignments, against 38% in the control group, although the difference was not significant, it does indicate that the 2% (not significant) lower completion rate for the game group was not due to fewer students making all assignments, but students obtaining a different grade.

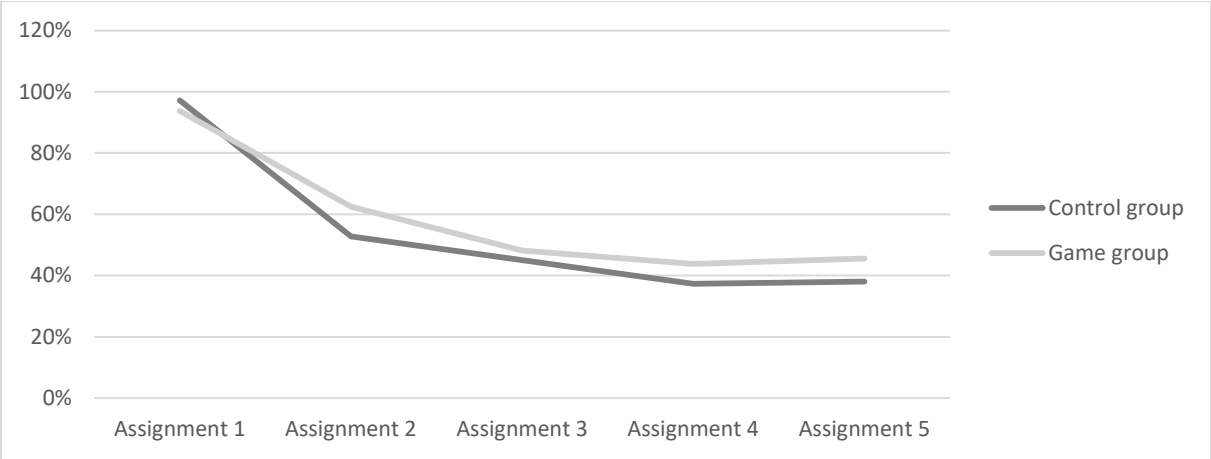


Figure 22: Assignments made per group, in percentage of total starters (N = 254)

Figure 22 shows that when looking at the percentage of students who start the course, the game group a higher percentage of students make all assignments in the game group. Note that both lines do not start at a 100%, as not all of the students who made at least one assignment (and were counted as *started)*, necessarily made the first assignment.

SG vs forum posts

No significant difference was found between the game group and the control group when comparing the average number of forums posts ($F = 0.447$; $t = 0.501$; $df = 252$; $p < 0.617$), see Table 27. In total there were 323 forum posts, 192 in the control group (N = 41) and 131 in the game group (N = 29), of which only 7 posts were about the game.

Table 27: Average number of forum posts per group

	Group	N	Mean	Std. Deviation	Std. Error Mean
Forum posts	Control group	142	1.35	2.888	0.242
	Game group	112	1.17	2.879	0.272

4.3.2 Supportive statistical analyses part 3

Age vs three retention variables

As in Sub-section 4.2.2 it was shown the two groups differed in the average age, the effect of age on the completion, assignments made and grade were analysed. Note that it could not be checked whether the age difference exists for the $N = 254$ students in the dataset used to analyse part 3, because this dataset (grade report and forum data) could not be coupled to the DelftX pre-questionnaire containing the student demographics. However, as the difference in age was significant when analysed using the questionnaire dataset (N = 90), the difference should be present in this dataset as well. Also, it was not possible to analyse the effect of age on the number of forum posts, as the forum posts dataset could not be coupled with the Delft pre-questionnaire dataset containing the student demographics. Completers were found to be 6 years younger than non-completers, 32 and 28 respectively ($p < 0.046$). Age was also significantly negatively correlated with the number of assignments made ($p<0.045$; $r = -0.213$) and the grade ($p < 0.033$; $r = -0.23$). The results of the analyses of the other students demographics on the completion and number of assignments made can be viewed in Appendix 18.

Serious game vs grade

After analysing the difference in grade between the serious game group and the control group, no significant difference was found. ($F = 0.069$; $t = -0.248$; $df = 252$; $p < 0.804$). However, looking at the grade for the students who completed the course, the game group (N = 31) scored significantly ($F = 2.542$; $t = -2.303$; $df = 71$; $p < 0.024$) higher (6 points) than the control group (N = 42). When looking at the grade for students who made all 5 assignments but did not complete the course, a difference was found between the game group and control group, but this difference was not, although almost, significant ($F = 5.296$; $t = 1.808$; adjusted $df = 21.6$; $p < 0.085$).

Game players per step

To get more insight into the lack of effect of the serious game on the completion, assignments made and forum posts, the serious game engagement was analysed further. It can be seen in Table 28 that of the students in the game group, only around 16% of the students who made the assignment at the end of the step, also finished one game. The distribution per assignment is visualized in Figure 23.

Table 28: Number of students who made an assignment and (not)played the corresponding game

Number of players	Number of players	Number of not-players	Total number of students	%
Made assignment 1	46	59	243	19%
Made assignment 2	26	44	145	18%
Made assignment 3	18	36	118	15%
Made assignment 4	15	34	102	15%
Made assignment 5	16	35	105	15%

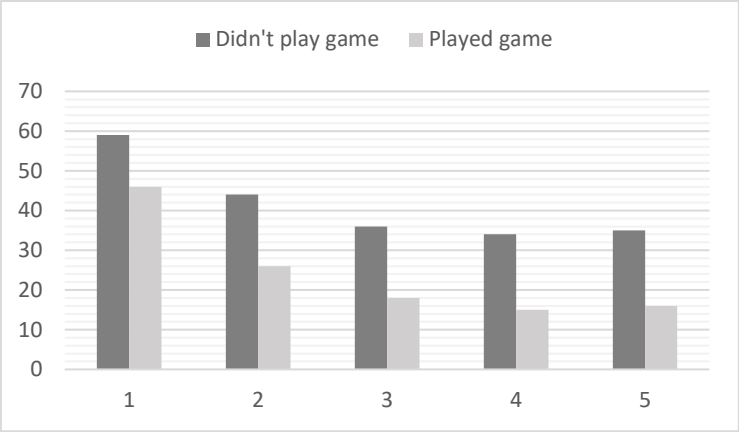


Figure 23: Students in the game group who made the assignment and did or did not play the game

4.3.3 Summary of findings theoretical framework evaluation part 3

This section presented the results of the serious game intervention on the completion, number of assignments made and number of forum posts, by analysing the difference between the game group and control group. When tested with contingency tables and independent sample-tests, no significant difference in completion, assignments made, or forum posts between groups was found, meaning none of the hypotheses are supported, as shown in Table 29.

Table 29: Results revised hypotheses theoretical framework part 3

H3.1	More students in the game group complete the course than students in the control group	Not supported
H3.2	Students in the serious game group make more assignments than students in the control group	Not supported
H3.3	Students in the serious game group make more forum posts than students in the control group	Not supported

To provide context for these results, several supportive analyses were conducted. First it was assessed if age was found to have a significant effect on completion and assignments made, as it was shown in Sub-section 4.2.2 that the students in the game group were on average 7 years older. The analysis showed that older students were less likely to complete, made fewer assignment and obtained lower grades.

When analysing the grades students in both groups obtained, no significant difference was found when all students in both groups were compared. However, when analysing the students who completed the course, it was found that students in the game group completed the course with a higher grade.

Looking closer at the engagement of student in the game group with the game, it was found that of the students who completed each step, approximately 16% played the game.

5. Discussion

After presenting the results of the experiment in the previous chapter, the results are discussed in this chapter for each part of the theoretical framework separately. In Section 5.1 the results regarding the relation between the factors and student retention are compared to the findings in literature study in Chapter 2. In Section 5.2 the effect of the serious game incorporation on the factors is discussed by reflecting back on the serious game design in Chapter 3. Lastly, the results of serious game incorporation on student retention in discussed by comparing the findings with other studies incorporation a serious game into a MOOC in Section 5.3.

5.1 Part 1: The factor effects on student retention

In Chapter 2, it was hypothesized that intrinsic motivation, extrinsic motivation, interactivity and perceived MOOC Quality influence student retention in MOOCs. In Chapter 4 it was shown that none of these hypotheses were supported. Due to the similarity between intrinsic motivation and extrinsic motivation regarding the literature study and results they are both discussed in Sub-section 5.1.1, while interactivity is separately discussed in Sub-section 5.1.2, and perceived MOOC quality in Sub-section 5.1.3.

5.1.1 Intrinsic and extrinsic motivation and completion

In the previous chapter it was shown that intrinsic motivation was not correlated with course completion.
It was hypothesized that intrinsic motivation would positively influence student retention as students often have intrinsic motivations to enrol for a MOOC (Gütl et al., 2014; Hew & Cheung, 2014; White et al., 2015; Xiong et al., 2015) and Xiong (2015) found, using a Structural Equation Modelling (SEM) approach, that both intrinsic and extrinsic motivations to enrol were a significant predictor for student engagement in a course. Furthermore, De Freitas et al. (2015) mention that a lack of motivation was a reason students gave to drop out of the course. It thus seems that the findings in this study contradict with findings in available literature. There are several reasons that could explain this contradiction.

De Freitas et al. (2015) did not specify whether their study concerned intrinsic, extrinsic, or another form of motivation, and she did not support this with data. Remarkably, it almost seems as if they assume that students are not motivated enough because lack of motivation is inherent to not completing a task.

There are several differences between the study of Xiong et al. (2015) and this study.

- The topic of the MOOC in their study was “Art”, which is quite different from the “analytical approach to solve complex problems” topic of the MOOC analysed in this study. Furthermore, the MOOC in this study was self-paced, whereas the MOOC studied by Xiong was not.
- They measure student retention as the number of days a student remains active in the course, while in this study it was not possible to look at students’ course engagement, but only at completion rates.
- Their study concerned a much larger sample (17.359 students) than the sample (52 students) in this study, which is especially important for a valid factor analysis.

- They measured intrinsic motivation with only one item in their SEM model; “I am taking this course out of general interest, curiosity, or enjoyment”, while in this study three separate items were used. Their items representing extrinsic motivation were similar to this study.
- In this research motivation was measured afterwards, by having students indicate with a 5-point Likert scale what motivated them to complete the course. This means it not the motivation when *enrolling*, but the motivation when *completing* was measured and used in the analysis. It was known in this study that 27% of the students enrolled for intrinsic reasons, and 73% for extrinsic reasons. However, as this variable was of dichotomous measurement level, it could only be analysed if students who are intrinsically motivated were more likely to complete than students who were extrinsically motivated, and not the individual contributions of intrinsic or extrinsic motivation to course completion.

5.1.2 Interactivity and completion

In the previous chapter it was shown that interactivity was not correlated with course completion in this study. Interactivity was incorporated into the theoretical framework as interaction with other students and teachers is a much mentioned reason for students to enrol for a course (Belanger & Thornton, 2007; White et al., 2015), the lack of interaction is also a much mentioned reason of students to drop out of a MOOC (De Freitas et al., 2015; Khalil & Ebner, 2014), and several authors found a significant correlation between interactivity and student retention (Adamopoulos, 2013; Hone & El Said, 2016).

The interactivity factor was represented in the questionnaire with items measuring to what extent students felt connected with their teachers, and what they thought on the amount of feedback they received during the course. As mentioned in Sub-section 3.3.4, the student-student interaction variable could not be taken into account in the factor analysis.

In their SEM model, Hone and El Said (2015) did find a significant correlation between student-teacher interaction and student retention. Their interaction operationalization is quite similar to this MOOC, but they did not just analyse if students completed the course or not, but how long they kept up with the MOOC, how many assignments they made and how much content they watched or read. As stated before, in this study these forms of course engagement could not be analysed. Furthermore, their post-questionnaire did not suffer a high non-response from students who did not complete the course.
Adamopoulos (2013) collected qualitative data concerning students’ review of MOOCs and used a sentiment analysis to construct quantitative data for several courses. This differed from the approach used in this study in which the interactivity factor was operationalized with items in a questionnaire. Furthermore, Adamopoulos measured retention by having students indicate themselves if they did not complete, partially completed or fully completed the course, while in this course student retention was based on data from the course platform itself.
Unfortunately, it was not possible to analyse the reasons for dropping out of this course and compare these findings with the findings of De Freitas et al. (2015) and Khalil and Ebner (2014) as of the 30 students who filled in the post-questionnaire, 10 students dropped-out of the MOOC, and of these 10 students only 2 indicated why they dropped out. This did not make it possible to see if the lack of interactivity was a reason to drop-out of the course, instead of only analysing if it was an important reason to complete it.

5.2.3 Perceived MOOC quality and completion

The perceived MOOC quality was not correlated with course completion in this study. Perceived MOOC quality was incorporated into the theoretical framework as Hew (2016) indicated that course resource diversity, active learning, and clear assignments with expositions all engage students in MOOCs. Furthermore, the literature study showed that student's sentiment for assignments and material (Adamopoulos, 2013), the perception on overall course design (Gütl et al., 2014) and perceived effectiveness of the MOOC (Hone & El Said, 2016), influence student retention.

The diversity of this term made it more difficult to operationalize the latent factor. The perceived MOOC quality factor score was represented in this study by students' perception on the amount of material and exercises available, and how challenging students found the course. It is therefore believed that operationalisation in this research is more closely related to that of Hew (2016) and Adamopolous (2013), than to that of Gütl (2014) and Hone and El Said (2016). Hew and Adamopolous specifically mention assignments, exercises, and course resources, while Gütl focused more on course design overall and Hone and El Said, focussed more on how effective students found the course. That being said, the DelftX post-questionnaire asked students to give an overall grade for the course, and also contained specific questions regarding the effectiveness of the course and both did not show significant differences regarding MOOC completion either. Lastly, in this study it was found that a student's perception of the quality of the MOOC is also based on how challenging the student found the course, which is not the case in any of the other studies.

5.2 Part 2: Discussing the impact of the serious game incorporation on the factors

In chapter 2, it was hypothesized that serious games could possible increase student retention by influencing the four identified factors: perceived MOOC Quality, interactivity, extrinsic motivation, and intrinsic motivation. As this was the first studies to research the influence of serious games on factors that were hypothesized to effect student retention in MOOCS, less literature is discussed than when discussing part 1. Instead, this section focusses more on how each factor was represented in the game.

5.2.1 Serious game and intrinsic motivation

Intrinsic motivation was negatively impacted by the serious game, according to the findings in Chapter 4. A surprising result as intrinsic motivation was included in the theoretical framework because serious games are heralded for their intrinsic motivational capabilities (Csikszentmihalyi, 1990; Garris et al., 2002; Thomas W. Malone, 1981).

The serious games were designed to be intrinsically motivating by representing the course content to make it *interesting*, setting a time constraint to be *challenging*, showing a leaderboard to make it *competitive*, talking about the game in the next step to raise students' *curiosity* and remembering students that the games were for *fun*. These are all important factors for intrinsically motivation instructional games according to Malone and Lepper (1987).

When analysing the quantitative questionnaire regarding the serious games, students indicated that the games were fun, challenging and curiosity raising, therewith contradicting the conclusion that the serious game caused the lower intrinsic motivation in the game group. However, when looking at the qualitative results, it was found that the games were not only mentioned as a positive aspect of the course, but that it was also often mentioned that the timing of the games needed to be improved. It could be that the students became frustrated or agitated from the challenge the time constraint introduced instead of motivated.

The finding that serious games negatively impact intrinsic motivation is different from the meta-analysis studies that did not find any impact of the serious game on intrinsic motivation (Wouters et al., 2013) and studies that found a positive impact on motivation (Clark et al., 2014). However, these studies were mostly in non-online learning environments. Wouters et al. give three possible explanations for serious games not influencing motivation: the absence of autonomy of when to play and what to play, the favouring of the instructional design over the entertainment design resulting in an instructive but less engaging game, and the possible ineffectiveness of measuring motivation with post-course questionnaires. The absence of autonomy was present in this MOOC, as students could decide for themselves if they wanted to play or not. It is likely that the instructional design was more present than the engaging design, due to the requirement that the game had to represent the course content. The possible ineffectiveness of measuring motivation with post-course questionnaire might be a reason, but in other studies this was not considered a problem (Adamopoulos, 2013; Hone & El Said, 2016).

5.2.2 The serious game and extrinsic motivation

It was concluded in the previous chapter that the serious game group had no different extrinsic motivation than the control group. This finding is less surprising as serious games are not necessarily known as methods to increase extrinsic motivation, and the relation between a serious game and extrinsic motivation was of more exploratory nature. The findings in this study at least indicate that the way the game was designed to increase extrinsic motivation was ineffective or too small to have a measurable impact.

The extrinsic motivational aspects of the game were remembering the students they were trying to become a creative problem solver, and making the game closely resemble applicability of the course contents with real-life examples, as most students indicated they wanted to apply the course content in their daily work. When analysing the quantitative questionnaire, the students indicated that they did find the game to remind them why they enrolled for this course, but not to great extent.

5.2.3 The serious game and interactivity

The students in the game group did not experience a higher interactivity than student in the control group. That a serious game might influence interactivity in a MOOC was based on the possibility games provide to interact with other students through competition or cooperation.

The interactivity was present in the serious game design by providing the possibility to compare the scores students achieved with other students via a leaderboard. Furthermore, the students were actively encouraged to discuss the game on the discussion forum. While the game was mostly focused on providing student-student interaction, the interactivity of students was mainly represented with answers to the statement “I felt connected with their teachers” and partly by the statement “I received enough feedback during the course”, which are both mostly concerned with student-teacher interaction and not student-student interaction. Furthermore, it was found that out of the 131 forums posts in the game group, only 7 concerned the game, leading to the conclusion that the students did not feel the need to discuss this serious game with other students.

5.2.4 The serious game and perceived MOOC quality

The students in the game group did not perceive the quality of the MOOC as higher than the students in the control group. That the serious game incorporation would influence the perceived MOOC quality was based on the possibility of increasing the course resource diversity, facilitate active learning and provide problem oriented assignments with clear expositions.

There is no question if there was more content available to the game group as there were a total of five games added to the course, and the control group did not get extra course material. The fact that this did not impact the perceived MOOC quality could indicate that the amount of course material and exercises were already enough, or that course resources diversity does not influence students’ perceptions on MOOC quality. Whether or not the games provided the opportunity for active learning and a more problem oriented approach with clear expositions in this MOOC is difficult to evaluate, as no extensive questionnaire was used to evaluate the game design.

5.3 Part 3: Discussing the impact of the serious game incorporation and student retention

In this section the results regarding the third part of the theoretical framework are discussed. When the theoretical framework was developed, it was mentioned that the third part of the framework represented the overall effect of the serious game on student retention, and could be used to measure if there was any effect on student retention to assure the serious game did not affect student retention via another way than the four identified variables.

The results in Chapter 4 showed no differences were found between the game group and the control group regarding the completion rate of both groups, the assignments made and the forum posts made. However, this result might have been influenced by the average student in the serious game group being 7 years older, and older students were less likely to complete, made fewer assignments and made fewer forum posts. Interestingly, the students in the serious game group who completed the MOOC, scored a higher grade.

These results are also not in line with the suggestions by Romero and Usart (2013) and Bernaert et al. (2015), who both found that students were very positive about the presence of the serious game and suggested the use of serious games to improve completion rates. Interesting differences between those studies and this study is that they evaluated the game with quantitative and qualitative questionnaires, but did not use a two-group experimental design. This study also evaluated the game with a questionnaire, but it was found that the questionnaire findings contradicted the findings from the group comparison. This indicates that the influence of the serious game should not only be measured by directly asking students what they thought of the game, as it might be that this does not provide accurate answers.

Another importance difference between these two studies and this study, is the presence of the serious game relative to other course content. In the study by Bernaert et al. (2015), the serious game was the main instruction method, while in the study of Romero and Usart (2013) the two serious games represented half of the MOOC. The presence of the games in this study was considerably smaller, especially with only about 16% of the students eventually playing the game in each step.

6. Conclusions, limitations and recommendations for future research

The purpose of this study was to gain insight into the possibilities serious games offer to increase student retention in MOOCs. It was decided that the most valuable insight would be obtained by first researching the theoretical extent to which serious games could influence student retention, and then test this theory by incorporating a serious game into a MOOC and conduct an experiment. This chapter describes the conclusions of our research regarding the main and sub-research questions in Section 6.1, before the scientific and societal contributions are discussed in Section 6.2 and 6.3. This thesis is concluded by presenting the limitations in Section 6.4 and recommendations for future research in Section 6.5.

6.1 Conclusions

The conclusions of this research are structured in accordance with sub-research questions 3, 4, and 5. First the conclusions regarding the sub-research questions are described separately, before the results are combined to answer the main research question.

Sub-research questions

RQ3: To what extent do extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality influence student retention in MOOCs?

While this study set out to analyse the effect of these four factors on student retention in MOOCs, it was not possible to analyse the effects of the factors on student engagement. It was only possible to conclude that the extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality did not have an effect on the course *completion* of students in this MOOC.

After comparing the findings of this study with other studies, it was found that most studies do use the engagement of students throughout the course to measure retention, which could explain why this study did not find any relation between the four factors and student retention. However, between those studies and this study were also several minor, and sometimes major, differences in how the factors were measured, how data was collected and what data analysis methods were used. Add to that that almost none of the MOOCs in the studies were self-paced, and it could be concluded that the findings of this study do not necessarily contradict those of other studies, as the diversity amongst MOOC studies is too high to state that the conditions in which the results were obtained are the same. Rather, these results add to a niche in the MOOC retention literature: the underlying factor structure influencing *completion* in *self-paced* MOOC.

RQ4: To what extent does a serious game affect extrinsic motivation, intrinsic motivation, interactivity, and perceived MOOC quality?

The results of this study showed that the serious game had a negative impact on the intrinsic motivation of students, and no impact on extrinsic motivation, interactivity, and perceived MOOC quality.

The negative impact of the serious game on intrinsic motivation was a surprising result, as serious games are theorized to be effective as instructive environment because of their capability to utilize the intrinsic motivational aspects of games. When analysing the quantitative game evaluation results, the games were regarded as a positive influence on all four factors, including intrinsic motivation. However, when looking at the qualitative evaluation of the game it was found that multiple students indicated that the timing of the games needed to be improved. It is therefore carefully concluded that the lower intrinsic motivation was caused by students disliking the timing in the games. This conclusion highlights the sensitivity of serious games to small design choices, as this small design choice had a large impact on the total game experience.

The lack of impact of the serious game on interactivity could be explained by interactivity in this study representing student-teacher interaction, while the game was designed to influence student-student interaction. Unfortunately, the student-student interaction could not be accounted for in this study as it proved not to be possible to include the measurement of student-student interaction in the factor analysis. However, with the interactivity factor not being influenced by the serious game, it can be concluded that this serious game could not be used to make students feel connected with their teachers.

The lack of impact of the serious game on extrinsic motivation and perceived MOOC quality is less surprising, as these two factors were less present in the serious game design than the other two factors. It is concluded that influencing students' extrinsic motivation via textual reminders of why they enrolled in a course did not have a measurable impact. Increasing students' perceived MOOC quality by providing an extra way of learning and increasing the course material with 5 small serious games also did also not appear to be possible with the serious game used in this research.

RQ5: Is there a difference in student retention when a serious game is incorporated into a MOOC?

This research question was answered by analysing the difference in course completion, assignment engagement and the number of forum posts. By being able to analyse both student completion and course engagement, both aspects of student retention as defined in this study were present. The results of these analyses show that the serious game incorporation did not affect course completion, the number of assignments made per student and the number of forum posts per student. However, the students in the game group were older, and higher age had a negative impact on the completion rate and number of assignments made, which might have distorted the effect of the serious game incorporation.

A further analysis of the serious game incorporation into the MOOC showed that on average only 16% of the students played the game, which might explain the low impact. It was not researched what made students decide to play the game or not. The lack of impact of the serious game is also not surprising considering the answer to the first research question. The game was specifically designed to influence the four factors, as these factors were found to influence student retention in related MOOC literature. With no relation between those factors and student course completion, it appears that even if the factors were influenced to great extent by the serious game, no effect regarding course completion would be obtained.

Main research question

In this thesis the impact of incorporating five small serious games into the Creative Problem Solving and Decision making MOOC was researched, to answer the main research question:

"To what extent can serious games be used to improve student retention in MOOCs?"

The findings in this research suggest that while the use of serious games is theoretically promising, their effectiveness is limited in practice.

This is not necessarily due to inability of serious games to motivate students, but due the limited understanding of factors affecting student retention in MOOCs. This was apparent in this study by the absence of the relation of intrinsic motivation, extrinsic motivation, interactivity, or perceived MOOC quality with student retention, while it was stated by several authors in different studies that these factors were found to be correlated with student retention. This led to the design of a serious game that aimed to impact these factors, but whatever impact there was did not lead to a change in student retention, as the factors turned out not to be correlated with student retention in this particular MOOC. It is therefore believed that in order for serious games to be able to improve student retention in MOOCs, first a better understanding of the low student retention is required.

6.2 Scientific contribution

This study contributes to scientific literature in several ways, mostly to MOOC literature regarding student retention and to serious game literature on designing games in online environments.

This study contributes to MOOC literature regarding student retention as it does not only build further upon previous findings, but actively tested if these findings hold their own in this study. By testing this in another MOOC it was shown that the results obtained in one MOOC cannot at all be assumed to be applicable to another MOOC, without a thorough comparison of both MOOCs. Furthermore, as the MOOC in this study was one of the few in available literature that was self-paced, this study specifically contributed by providing insights into factors affecting completion rates in self-paced MOOCs.

This study also contributed to the debate of how MOOC completion rates should be measured, by arguing that the students should only be counted as started when they have made notable progress in the course. This differs from counting enrolled students, or counting students who showed any indication of activity as started, and is considered fair as it allows for better interpretation of a completion rate of a course. It is more valuable to known how many students completed when it concerns students who actually showed an intention to complete it.

By exploring the possibilities serious games offer to motivate students in online learning, a contribution was made to literature regarding serious games. The results of this study contribute to the debate on the motivational impact of serious games, by showing that the impact of the serious game in this study led to lower intrinsic motivation. This is a valuable contribution, as several authors suggest the use of serious games to improve students' motivation in MOOCS, while there is limited research available on the effects of serious games in online learning.

This study also contributed to serious gaming literature and online learning both, by being the first to study the effects of a serious game incorporation with a true experimental design. None of few studies that incorporated a serious game into a MOOC have used a true experimental design, making it often difficult to imply causation between the serious game integration and the achieved results.

Lastly, this study contributed to serious gaming literature by not only being one of the first true experimental design studies regarding serious games in online learning, but also showing that the quantitative game evaluation questionnaire results contradicted the results obtained when comparing the serious game group with the control group. This suggests that the influence of the serious game should not only be measured by directly asking students what they thought of the game, as it this does not always provide accurate results.

6.3 Societal contribution

In the problem exploration it was stated that the low student retention in MOOCs leads to valuable resources being ineffectively spent to achieve university goals, less income being generated and university level education being less effectively distributed around the globe. Although this research does not provide a solution to this problem, it did contribute valuable knowledge.

This study showed the practical requirements and design considerations in order to successfully integrate a serious game into edX. It showed that the Iframes edX provides to enable the integration of interactive elements could also be used integrate a serious game with relative ease, and that by using the Iframe it was possible to provide an anonymous user-id of the students with which it is possible to collect individual game data. These findings could be used by MOOC facilitators to explore the use of interactive elements in their courses and improve their courses overall.

It was also shown that careful design considerations have to be made when a serious game is chosen as method to solve this particular problem, as a small wrong design choice could impact the motivation of a student negatively. The most prominent example of this is the negative findings of this study regarding intrinsic motivation. These findings can be used to decide to refrain from using strict time restrictions in future serious game designs and

therewith preventing having the same negative impact found in this study. It is believed that small steps like these would eventually lead to improved serious game designs to increase student retention in MOOCs.

Lastly, this study showed that edX provides the possibility to completely separate two groups of students with relative ease, and randomly allocate students to both groups, which could lead to an experimental environment with high internal validity. However, researchers and MOOC facilitators should be careful with assuming that the groups are equal, as the supposed random allocation of students by edX resulted in a different number of students per group, and a significant difference in age between the students.

6.4 Limitations

The findings of this study are limited by several factors, this section describes those limitations and links them to the specific findings that are implicated by them.

External validity limitations

The external validity of the results to other MOOCs is low as only one MOOC with a tailor made serious game was researched. This means that the findings of this study can only be generalized to MOOCs with similar course design, topic and type of enrolled students. This would be MOOCs that are self-paced, take approximately 20 to 40 hours to complete and teach an analytic approach or at least concern an analytical or technical topic, with students who are mostly working professionals with an average age of 36 years. The low external validity does also applies to the serious game used in this MOOC. As this serious game was specifically designed to increase student retention in this MOOC, the negative effect of the serious game on intrinsic motivation cannot be generalized to serious games in general.

Measurement instruments limitations

There were several limitations regarding the measurement instruments used in this study. First, due to the presence of several questionnaires from DelftX there was little room available to add an additional questionnaire to measure the four factors from the theoretical framework and for evaluating the serious game. Meaning only a limited number of questions could be used to represent the construct-item relation used in the factor analysis and a limited number of questions could be used evaluate the serious game experience of students. Second, there was a limited response for the pre-questionnaire (97), mid-questionnaire (45) post-questionnaire (31) and additional questionnaire (52) used in this research due to the limited time available to collect data and the high non-response of students. Third, in the additional questionnaire a mistake was made that replaced the highest Likert-scale indication “Strongly agree” with “Somewhat agree”. Although this was the case for both groups, it might still have had an impact as the one of the groups could have wanted to give a higher score on one of the items, possibly making the difference between the groups regarding the factor scores smaller. Fourth, an analysis of the respondents showed that most students who filled in the questionnaires completed the MOOC, indicating a self-selection bias. This self-selection bias meant that results regarding the student engagement in the course could only be somewhat generalized to students who completed the course. Meaning that the results of theoretical framework evaluation part 2 and 3 must be interpreted as: “the serious game incorporation had no effect on the four factors, the number of assignments made and number of forum posts, *for students who completed the course*”.

Factor analysis limitations

The impact of the low response to the additional questionnaire would be limited, if the responses did not have to be used to measure latent variables. For this a factor analysis was required, and for a factor analysis a minimum of N = 100 is advised. Furthermore, in the factor analysis the factors were represented with only 2 or 3 items, which is the minimum number of items per factor according to Matsunaga (2010). The interactivity factor was represented by two items, of which one had twice as much influence on the factor score as the other. Although there was a simple structure obtained with the factor analyses, indicating that the questionnaire did indeed measure the four latent factors, the validity of the factor analysis is questionable.

Data limitations

Several issues surfaced regarding the data used for this research. First, it was not found to be possible to gather data on the number of videos watch per student. Second, it was not possible to couple the dataset containing the serious game engagement and number of forum posts per student. Third, due to the self-selection bias, almost all students who filled in the additional questionnaire made all 5 assignments, making it not possible to analyse the effect of the four factors on the number of assignments made. These three limitations did not make it possible to analyse the effect of the four factors on all three variables commonly used to represent student course engagement: the number of videos watched, the number of assignments made and the number of forum posts.

Serious game limitations

The serious game design and development was also limited in development time and budget, making it only possible to design a serious game that used existing game mechanisms at InThere, of a serious game that was initially designed for another purpose. This meant some of the factors were only tried to influence with textual encouragement, instead of game mechanics, possibly reducing the effectiveness of the serious game.

6.5 Recommendations for future research

There are several recommendations made for future research: (1) rerun this experiment with improved conditions, (2) analysing underlying factor structure influencing student retention in MOOCs, (3) using similar serious game in a different MOOC, (4) using different serious game in a similar MOOC.

1. Rerun this experiment with improved conditions

The first recommendations for future research would be to reproduce the findings of this thesis taking into account the limitations presented in the previous section. The easy improvements to be made would be to add the “strongly agree” option in the additional questionnaire and increase the time limit for each serious game. The research quality could be improved further by a significant increase in the number of respondents. Furthermore, it would be recommended to more actively pursue students who dropped out of the MOOCs to fill in the questionnaire, to prevent collecting responses from only students who complete the course. Lastly, expanding the data on student engagement with the number of videos watched or the number of non-obligatory assignments made would offer valuable insight in the effect of the four factors into the student engagement throughout the course, instead of only course completion, which could then also be used to better compare this study to other studies analysing student retention.

2. Analysing underlying factor structure influencing student retention in MOOCs

In the conclusion of this research it was stated that in order to better match the possibilities serious games offer as solution to the problem MOOCs currently suffer from, it is first needed to further analyse the cause underlying the low student retention in MOOCs. Despite several authors finding a relation between student retention and extrinsic motivation, intrinsic motivation, interactivity and perceived MOOC quality, these findings could not be reproduced. Whether this was due to the low validity of the factor analysis, the small sample size, or the different nature of the MOOC and its students analysed in this study is not known. It would therefore be interesting to conduct a Confirmatory Factor Analysis based on the Explorative Factor Analysis in this research, on a different dataset. However, as the factor analysis in this research was constructed with a non-validated questionnaire that was limited in the number of items per construct, it would also be interesting to construct a new questionnaire based on the questionnaire on this research and similar questionnaires from other studies, perform an Exploratory Factor Analysis to identify how the underlying latent factors are related, and then confirm these relations with a Confirmatory Factor Analysis on not one, but several MOOCs. This would result in a better understanding on what influences student retention in MOOCs in general, and prevent future studies like these having to first analyse extensively what influences student retention in their respective MOOC, before a solution can be incorporated.

3. Using a similar serious game in different MOOC

It would be interesting to see if the same or a similar serious game would have the same effects in a different MOOC than the MOOC used in this study, for example a MOOC that is not self-paced, has a different topic, or is followed by different students. This would provide valuable insight into the effectiveness of the serious game in varying conditions, and show which type of MOOCs could benefit from this type of serious game most. It would also show if the serious game was not suitable to effect student retention in the particular MOOC used in this study, or that the serious game was not suitable to effect student retention at all.

4. Using a different serious game in a similar MOOC

Studying if a different serious game would have effects in this MOOC could provide valuable insights into what kind of serious game this type of MOOC would benefit from most. As in this study out of the 6135 students who enrolled for the course, only 254 finished one assignment, it could be interesting to research the possibilities serious games offer to persuade more students to start the course. It was not possible to develop and incorporate a serious game in which students or teacher could interact with each other, due to the increased game complexity and limited budget and time available to this research. However, with many authors indicating that the lack of interaction is an important factor influencing student retention, further research on to what extent a serious game focused on creating interaction between students and teachers could influence student retention is suggested. Lastly, seeing that on average only 15% of the students played a game in the game group in this study, it is suggested to research what would happen if the game was to be made obligatory, therewith increasing the presence of the serious game in students' MOOC experience.

Bibliography

Abt, C. C. (1977). *Serious games*. New York: Viking Press. https://doi.org/10.1145/2465085.2465091

Adamopoulos, P. (2013). What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. In *Thirty Fourth International Conference on Information Systems* (Vol. 49). New York.

Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences* (4th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Baarda, B. (2006). *Basisboek Methoden en Technieken* (4th ed.). Houten: Noordhoff Uitgevers.

Belanger, Y., & Thornton, J. (2007). Bioelectricity: A Quantitative Approach. Boston, MA: Springer US.

Berge, Z. L., & Huang, Y.-P. (2004). A model for sustainable student retention: A holistic perspective on the student dropout problem with special attention to e-Learning. *Deosnews*, *13*(5), 1–26.

Bernaert, O., Crepon, R., & Dhorne, L. (2015). MOOC and serious game an educational approach on transfer and action. Proceedings of the 43rd SEFI Annual Conference 2015 - Diversity in Engineering Education: An Opportunity to Face the New Trends of Engineering, SEFI 2015, 8. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84968831205&partnerID=40&md5=20dcc5ffeca22f9a5f9858926875e534

Blanco, A., Serrano, A., Martinez, I., & Fernandez-Manjon, B. (2013). Integrating serious games into e-learning platforms: Present and future. In I. A. Stanescu (Ed.), *The 9th International Scientific Conference eLearning and software for Education* (pp. 380–386). https://doi.org/10.12753/2066-026X-13-145

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment*, *8*, 13–25. Retrieved from papers3://publication/uuid/D2DE896D-3743-46EA-820F-2ACBACC1A78E

Clark, D. B., Tanner-smith, E. E., & Killingsworth, S. (2014). *Digital games, design, and learning: A systematic review and meta-analysis (executive summary)*. Menlo Park. Retrieved from https://www.sri.com/sites/default/files/publications/digital-games-design-and-learning-executive_summary.pdf

Clow, D. (2013). MOOCs and the funnel of participation. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 185–189. https://doi.org/10.1145/2460296.2460332

Cooper S, F, K., A, T., J, B., J, L., M, B., … Z., P. (2010). Predicting protein structures with a multiplayer online game. *Nature*, *466*(7307), 756–760. https://doi.org/10.1038/nature09304

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper & Row. Retrieved from http://www.bates.edu/purposeful-work/files/2015/03/Csikszenthmihalyi-1990.pdf

Davis, H., Dickens, K., Leon, M., Del Mar Sanchéz Vera, M., & White, S. (2014). MOOCs for universities and learners: An analysis of motivating factors. *CSEDU 2014 - Proceedings of the 6th International Conference on Computer Supported Education*, *1*, 105–116. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84902320154&partnerID=40&md5=08c13effa36f8b51cdc39f9cebe8cfd2

De Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology*, *46*(3), 455–471. https://doi.org/10.1111/bjet.12268

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, 9. https://doi.org/10.1145/2181037.2181040

Dharwal Shah. (2016). Monetization Over massiveness: Breaking down MOOCs by the numbers in 2016. Retrieved January 13, 2017, from https://www.edsurge.com/news/2016-12-29-monetization-over-massiveness-breaking-down-moocs-by-the-numbers-in-2016

Duke, R. D., & Geurts, J. (2004). *Policy Games for Strategic Management*. Rozenberg Publishers. Retrieved from http://books.google.com/books?id=XGUdoRPFx30C&pgis=1

Dym, C. L., Little, P., Orwin, E. J., & Spjut, E. (2009). *Engineering Design: A Project-based introduction*. John Wiley and sons. https://doi.org/10.1017/CBO9781107415324.004

Dynamic Systems Development Method ltd. (2014). MoSCoW Prioritisation. Retrieved September 8, 2017, from https://www.dsdm.org/content/moscow-prioritisation

Factor Analysis | SPSS Annotated Output. (2017). Retrieved September 2, 2017, from https://stats.idre.ucla.edu/spss/output/factor-analysis/

Freire, M., Del Blanco, Á., & Fernańdez-Manjón, B. (2014). Serious games as edX MOOC activities. In *Global Engineering Education Conference (EDUCON), 2014 IEEE* (pp. 867–871). https://doi.org/10.1109/EDUCON.2014.6826198

Freire, M., Martínez-Ortiz, I., Moreno-Ger, P., & Fernández-Manjón, B. (2015). Requirements for educational games in MOOCs. In *Global Engineering Education Conference, (EDUCON), 2015 IEEE* (pp. 993–997). https://doi.org/10.1109/EDUCON.2015.7096094

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*(4), 441–467. https://doi.org/10.1177/1046878102238607

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a Massive Open Online Course. *American Educational Research Journal*, *52*(5), 925–955. https://doi.org/10.3102/0002831215584621

Gütl, C., Rizzardini, R. H., Chang, V., & Morales, M. (2014). Attrition in MOOC: Lessons learned from drop-out students. *Communications in Computer and Information Science*, *446 CCIS*, 37–48. https://doi.org/10.1007/978-3-319-10671-7_4

Harteveld, C. (2011). *Triadic Game Design: Balancing Reality, Meaning and Play*. Springer Science & Business Media. https://doi.org/10.1007/978-1-84996-157-8

Hatton, S. (2008). Choosing the "right" prioritisation method. In *Proceedings of the Australian Software Engineering Conference, ASWEC* (pp. 517–526). https://doi.org/10.1109/ASWEC.2008.4483241

Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, *47*(2), 320–341. https://doi.org/10.1111/bjet.12235

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, *12*, 45–58. https://doi.org/10.1016/j.edurev.2014.05.001

Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers and Education*, *98*, 157–168. https://doi.org/10.1016/j.compedu.2016.03.016

Jansen, D., & Schuwer, R. (2015). *Institutional MOOC strategies in Europe: Status report based on a mapping survey conducted in October - December 2014*. Heerlen. Retrieved from https://www.surfspace.nl/media/bijlagen/artikel-1763-22974efd1d43f52aa98e0ba04f14c9f3.pdf

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *International Review of Research in Open and Distance Learning*, *15*(1), 133–160. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84894521471&partnerID=40&md5=fd6b392a7395548d15bb8021b875e97a

Khalil, H., & Ebner, M. (2014). MOOCs completion rates and possible methods to Iiprove retention-A literature review. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1234–1244). Chesapeake.

Kim, T. D., Yang, M. Y., Bae, J., Min, B. A., Lee, I., & Kim, J. (2016). Escape from infinite freedom: Effects of constraining user freedom on the prevention of dropout in an online learning context. *Computers in Human Behavior*, *66*, 217–231. https://doi.org/10.1016/j.chb.2016.09.019

Kirk, R. E., Kirk, & E., R. (2003). Experimental Design. In *Handbook of Psychology*. John Wiley and sons. https://doi.org/10.1002/0471264385.wei0201

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13* (pp. 170–179). https://doi.org/10.1145/2460296.2460330

Koller, D., Ng, A., & Chen, Z. (2013). Retention and intention in Massive Open Online Courses: In depth. Retrieved May 31, 2017, from http://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses-in-depth

Leire, C., McCormick, K., Richter, J. L., Arnfalk, P., & Rodhe, H. (2016). Online teaching going massive: Input and outcomes. *Journal of Cleaner Production*, *123*, 230–233. https://doi.org/10.1016/j.jclepro.2015.12.014

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: a systematic study of the published literature 2008-2012. *International Review of Research in Open and Distance Learning*, *14*(3), 202–227. https://doi.org/10.3329/bjms.v12i4.16658

Maalej, W., Pernelle, P., Marty, J.-C., & Carron, T. (2015). Integration of game based learning into a TEL platform: Application to MOOCs. In R. Munkvold & L. Kolas (Eds.), *Proceedings of the 9th European Conference on Games-based Learning* (pp. 321–330). Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84955079483&partnerID=40&md5=c223ed37e3d0687c7f055eb87294a524

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science*, *5*(4), 333–369. https://doi.org/10.1016/S0364-0213(81)80017-1

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude learning and instruction* (Vol. 3, pp. 223–253). Hillsdale, New Jersey: Lawrence Erlbaum Associates. https://doi.org/10.1016/S0037-6337(09)70509-1

Martí-Parreño, J., Méndez-Ibáñez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: a bibliometric and text mining analysis. *Journal of Computer Assisted Learning*, *32*(6), 663–676. https://doi.org/10.1111/jcal.12161

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, *3*(1), 97–110. https://doi.org/10.4090/juee.2008.v2n2.033040

Meyer, K. A. (2014). Student engagement in Online Learning: What works and why. *ASHE Higher Education Report*, *40*(6), 1–114. https://doi.org/10.1002/aehe.20018

Michael, D. R., & Chen, S. L. (2005). *Serious Games: Games That Educate, Train, and Inform*. *Education*. Boston: Thomson Course Technology PTR. https://doi.org/10.1021/la104669k

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). E-Learning, online learning, and distance learning environments: Are they the same? *Internet and Higher Education*, *14*(2), 129–135. https://doi.org/10.1016/j.iheduc.2010.10.001

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in Massive Open Online Courses. *Educational Researcher*, *43*(9), 421–432. https://doi.org/10.3102/0013189X14562423

Petri, G., Gresse, C., & Wangenheim, V. (2016). How to evaluate educational games: A systematic literature review. *Journal of Universal Computer Science*, *22*(7), 992–1021.

Rice, J. W. (2012). The gamification of learning and instruction. *International Journal of Gaming and Computer-Mediated Simulations*, *4*(4), 81–83. https://doi.org/10.4018/jgcms.2012100106

Romero, M., & Usart, M. (2013). Serious games integration in an entrepreneurship Massive Online Open Course (MOOC). In *International Conference on Serious Games Development and Applications* (pp. 212–225). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-40790-1_21

Rouse, M. (2015). What is IFrame (Inline Frame)? - Definition from WhatIs.com. Retrieved September 4, 2017, from http://whatis.techtarget.com/definition/IFrame-Inline-Frame

Salen, K., & Zimmerman, E. (2004). *Rules of Play - Game Design Fundamentals*. MIT Press Cambride. https://doi.org/10.1093/intimm/dxs150

Saraguro-Bravo, R. A., Jara-Roa, D. I., & Agila-Palacios, M. (2016). Techno-instructional application in a MOOC designed with gamification techniques. In *2016 3rd International Conference on eDemocracy and eGovernment, ICEDEG 2016* (pp. 176–179). https://doi.org/10.1109/ICEDEG.2016.7461717

Simões, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior*, *29*(2), 345–353. https://doi.org/10.1016/j.chb.2012.06.007

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, *64*(2), 489–528. https://doi.org/10.1111/j.1744-6570.2011.01190.x

Steiger, J. H., & Schönemann, P. H. (1978). A history of factor indeterminacy. In *Theory Construction and Data Analysis in the Behavioral Sciences* (pp. 136–178). San Francisco: Jossey-Bass.

Susi, T., Johannesson, M., & Backlund, P. (2007). Serious games – An overview. *Elearning*, *73*(10), 28. https://doi.org/10.1.1.105.7828

Tan, C. T. (2013). Towards a MOOC game. In *Proceedings of The 9th Australasian Conference on Interactive Entertainment Matters of Life and Death - IE '13* (pp. 1–4). New York, New York, USA: ACM Press. https://doi.org/10.1145/2513002.2513040

Vaibhav, A., & Gupta, P. (2014). Gamification of MOOCs for increasing user engagement. In *2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE)* (pp. 290–295). IEEE. https://doi.org/10.1109/MITE.2014.7020290

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013-2015. *International Review of Research in Open and Distance Learning*, *17*(2), 198–221. https://doi.org/10.19173/irrodl.v17i2.2448

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research*, *34*(3), 229–243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM

White, S., Davis, H., Dickens, K., León, M., & Sánchez-Vera, M. M. (2015). MOOCs: What motivates the producers and participants? In S. Zvacek, M. T. Restivo, J. Uhomoibhi, & M. Helfert (Eds.), *Communications in Computer and Information Science* (Vol. 510, pp. 99–114). Barcelona: Springer International Publishing. https://doi.org/10.1007/978-3-319-25768-6_7

Wilkowski, J., Deutsch, A., & Russell, D. (2014). Student skill and goal achievement in the mapping with google MOOC. In *ACM conference on Learning @ scale conference* (pp. 3–10). New York. https://doi.org/10.1145/2556325.2566240

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Xiong, Y., Li, H., Kornhaber, M. L., Suen, H. K., Pursel, B., & Goins, D. D. (2015). Examining the relations among student motivation, engagement, and retention in a MOOC: A structural equation modeling approach. *Global Education Review*, *2*(3), 23–33. Retrieved from http://ger.mercy.edu/index.php/ger/article/view/124

Appendices

Appendix 1: Description of MoSCoW method

In Sub-section 3.2.1 the MoSCoW method is used to describe the design space for the serious game. In this appendix a brief description of this method and its pro's and con's is given. The **M**o**SC**o**W** is an abbreviation of the following:

- **M**ust have
  The requirements that a project has to deliver in order for it to be called a success. If one of these requirements is not met, the project is considered a failure. Must have's take up to a maximum of 60% of the requirements. The other 40% being divided by should have's and could have's.

- **S**hould have
  Should have's are used to describe component that could give large problems when not incorporated, but are not vital to the project. It often means someway of workaround is possible.

- **C**ould have
  The differentiation between could have and should have is less strict than between should have and must have. Could have components are the components that are the first to be dropped when a lack of time or money occurs or if the quality standard is lowered.

- **W**on't have
  Won't have are the requirements of which the developers have agreed they will not be met in this project. Describing Won't have's help with showing the demarcation of the project and could serve as fields for future development or research.

The MoSCoW method has benefits over the use of high, medium or low prioritizing as the categories do not have to be defined, it is clear what can be expected and there is no "vague" middle option (Dynamic Systems Development Method ltd, 2014). It is preferred over sequential prioritizing (e.g. 1, 2, 3, 4) as it does not give rise to conflicts regarding requirements of similar importance (Dynamic Systems Development Method ltd, 2014). According to Hatton (2008), who compared several prioritisation methods, benefits of the MoSCoW method are that is quick and easy to perform, which fits nicely with the limited timeframe for this research; it is possible to add requirements later, which fits with the iterative and dynamic design approach of TGD; and it is suitable for goals and high level requirements. She notes that a downside is that it is not suitable to list more detailed requirements and rank this information accordingly, but since the serious game developed can be considered a relatively small project, this is not deemed a problem.

Appendix 2: CSPD MOOC 2016 analysis

The course from the previous years was evaluated using the data from the edX platform and the questionnaires, provided by DelftX. These results can be used to identify weak points or strong points the serious game design could benefit from. The analyses was both quantitative and qualitative.

Summary of findings quantitative analysis

This section is summarized in 30 below. From the data from the edX-platform is could be derived that the number of student that attempted each multiple choice assignment, drops significantly after the first two assignments (1279, 743, 539, 467, 404) and seems to follow the "funnel of participation's" characteristic drop (Clow, 2013). The student activity in forum posts dropped after the start in June, to remain stable till December. Only 5% of the students started more than 50% of the videos, 1% started more than 75% of the videos.

In the pre-questionnaire (N = 1700-1900) it was indicated that the large majority of students were working professionals (65%), followed by students (16%), job-seekers (9%) and recently graduated (5%). The main reason for enrolment was "to use what they will learn in their daily work" (50%), interest in the subject (24%) and to improve future career prospects (23%). In the mid-questionnaire *(N=230),* halfway through the course, students gave their "happiness" a 4 out of 5.

In the post-questionnaire quantitative (N = 145-160) results the student gave the course overall a 7.18 out of 10. Students felt the course was "about right" regarding difficulty, workload and duration. Students rated activity of lecturers as neutral to slightly positive, even though they did not participate in forums that much. Interaction with other students, helping other students with their coursework, receiving help, communicating with the instructor/assistants and using the course to network were rated neutrally. Students rated they spent about as many hours as was expected. Students felt the course had value because they learned a lot and could apply it in their daily work.

Table 30: Summarized findings of CSPD MOOC 2016

Questionnaire		**Positive**
Data edX		Funnel of participation is observed regarding assignments.
		Forum activity quickly drops after start, then remains stable
Pre-questionnaire	Closed	65% working professionals, 16% students, 9% job-seekers, 5% recently graduated.
		50% can use course in daily work, 25% interest and 23% to improve career.
Mid-questionnaire	Closed	Happiness was rated 4 out of 5
Post-questionnaire	Closed	Course scored 7.18 out of 10
		Teachers were regarded slightly positive
		Value of the course was that it learns a lot and is applicable in daily work
	Open	Mentioned positive
		General content of the course, approach, methods and instruction, and examples
		Mentioned negative
		Methods, examples and assessment

Summary of findings Qualitative analysis

The qualitative analysis was conducted by structuring the answers to the open questions by topic and by being negative or positive. The open questions can be found in Appendices 3, 4, and 5 containing the DelftX questionnaires. In the qualitative (N= 4-90) results it could be found that the general content of the course, approach, methods and instruction, and examples are mentioned the most as positive aspects of the course. The methods, examples and assessment are mentioned the most as negative aspects of the course. Interestingly, the examples and methods are thus both often mentioned as negative (in particular scenario analysis) and positive.

Table 31: Qualitative CSPD MOOC 2016 evaluation

	Number responses	Positive		Negative	
Content general	48	33	69%	15	31%
Tools	80	57	71%	23	29%
Approach	29	29	100%	0	0%
Instruction general	36	25	69%	11	31%
Instructors	3	3	100%	0	0%
Videos	20	12	60%	8	40%
Material	14	10	71%	4	29%
Examples	45	24	53%	21	47%
Applicability	12	6	50%	6	50%
Forum	13	4	31%	9	69%
Assesment	29	9	31%	20	69%
Other	17	10	59%	7	41%
	346	**222**	64%	**124**	36%

Conclusion

From this analysis it can be concluded that most participants are working professional that seek to improve their career or want to at least apply it in their daily work. This is taken into account in the serious game design by ensuring the game does not come across as childish and give the participants the feeling they are not taken seriously. Furthermore, it is important to emphasize with the game that the course is indeed applicable to their daily life. From the qualitative results it can be concluded that the general approach highly valued and the methods are often specifically mentioned, albeit in a negative or positive way. To emphasise this course's apparent strong point, the methods and the general approach should be prominent in the game design.

Appendix 3: DelftX pre-questionnaire

We start this survey with some questions about why you chose to enrol in this course and your prior knowledge in this particular field of expertise.

1. How did you discover this course? I found out about this course through:
    - A topic search on Google (or any other search engine)
    - The TU Delft Online Learning webpage
    - Social media (e.g. LinkedIn, Twitter, Facebook, etc.)
    - A newspaper or magazine I've read
    - Someone I know told me about this course
    - The Delft University of Technology homepage
    - An online article I've read
    - A flyer or poster I've seen
    - An online advertisement
    - Some other way …

2. Why did you enroll in this course? Choose the answer that best describes your motivation:
    - I can use what I will learn in this course in my daily work activities
    - I am taking this course to improve my future career prospects
    - I want to get a degree in this field (or related area)
    - I am interested in this subject and want to learn more about it
    - I know the instructors of this course and am very interested in their work
    - Other …

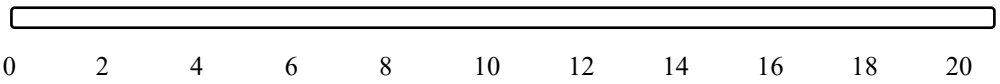
3. How important were the following factors in your decision to enroll in this course?
    - A recommendation from someone else
    - The relevance of the course topic for my occupation or work
    - My general experience with online university courses
    - The status and reputation of Delft University of Technology

4. How familiar were you with Delft University of Technology prior to this course? Choose the answer that best describes your familiarity with Delft University of Technology:
    - Not at all familiar
    - Slightly familiar
    - Somewhat familiar
    - Moderately familiar
    - Very familiar

The next questions are about your experience(s) with online learning, which course aspects you find important, and how comfortable you are communicating in English.

5. How many online courses have you taken before?
    - None, I have not taken an online course before
    - 1
    - 2
    - 3
    - 4
    - 5 or more

6. On average, how many hours per week can you dedicate to this course?



7. How do you intend to participate in this course? Choose the answer that most closely describes your intentions:
    - I want to complete this course
    - I will only do the specific parts of this course that interest me
    - I am just checking this course out
    - Other …

8. Which activities do you plan on doing in this course? Will you ...
    - Do assignments and quizzes
    - Participate in the forums
    - Watch the videos
    - Read course materials

How do you prefer to learn? Please select the option below that best indicates your preference, for each pair of statements. The middle button means that you like both ways of learning the same amount.

9. How do you prefer to learn?
    - Learn on your own with low or inexistent support from the course team and peers
    - Both
    - Learn with the support of the course team and peers (community) and contribute to other's learning

10. How do you prefer to learn?
    - Set your own pace (e.g. self-paced course where you set your deadlines)
    - Both
    - Let the pace be set by the course team (e.g. comply with deadlines)

11. How do you prefer to learn?
    - Learn by watching videos, reading and answering quizzes
    - Both
    - Learn by solving problems and doing assignments

12. How do you prefer to learn?
    - Make your own choices (e.g. work individually or in group, select study topics)
    - Both
    - Work with course materials and learning activities set by the course team

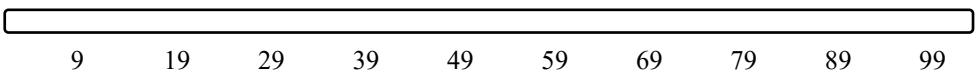
13. How do you prefer to learn?
    - Work with theoretical examples
    - Both
    - Work with real world examples

14. How would you rate your level of English proficiency?
    - Poor
    - Fair
    - Average
    - Good
    - Very Good

15. How comfortable are you communicating in English?
    - Very uncomfortable
    - Uncomfortable
    - Neither comfortable nor uncomfortable
    - Comfortable
    - Very comfortable

We conclude this survey with some questions about your personal situation and background.

16. What is your age?



17. What is your gender?
    - Male
    - Female
    - Other

18. What is your (first) nationality? Please select the corresponding continent and country.
    - Continent (*6 choices*)
    - Country (*63 choices*)

19. Do you hold a dual nationality?
    - Yes
    - No

20. What is the highest degree or level of education you have completed?
    - Elementary school degree
    - Secondary school (e.g. high school or equivalent) diploma / degree
    - College / Associate's degree
    - Professional / Vocational degree
    - Bachelor's degree
    - Master's degree
    - Postgraduate / Graduate degree
    - Ph.D.
    - Other …

21. Which of the following best describes your current situation?
    - Student
    - Recently graduated
    - Working professional
    - Looking for a job
    - Parent / care-giver
    - Retired
    - Other …

22. We are doing a lot of exciting research at Delft University of Technology. Would you be willing to receive information about participation in one of our future research projects? This could be a survey, interview, or something entirely different (e.g. building a weather balloon)! Let us know by clicking 'Yes' below, and you might receive an email from us in the future.
    - Yes
    - No

    We thank you for your time spent taking this survey.

Appendix 4: DelftX mid-questionnaire

The course has been running for several weeks now, and we're very interested to know what you think about it so far! Could you spare us a minute to give some feedback? Great, thanks!

1. Please let us know how much you're enjoying the course through the smiley face (click on the smiley to activate it).

2. Which aspect of this course do you especially like? Please name one.

   ……

   ……

3. Which aspect of this course would you like us to change or improve? Please name one.

   ……

   ……

Appendix 5: DelftX post-questionnaire

1. Since the start of the course, how would you describe your participation level?
   In order for the survey to unfold correctly you need to answer this question, but don't worry, all the other questions are optional!
   - o I did not participate in this course at all
   - o I did not participate in this course, but I did browse around a little bit (→A)
   - o I only looked at specific parts of the course that I was interested in (→B)
   - o I participated in the course, but I stopped participating along the way (→C)
   - o I participated in the course up until the end

   A. Could you please describe the reason(s) why you did not start the course?

   ……
   ……

   B. Could you please describe which specific parts of the course you were interested in and why?

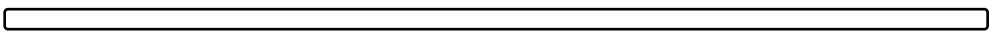
   ……
   …… …...

   C. Could you please describe the reason(s) why you decided not to continue with the course? Choose the one that applies the most.

   ……
   ……

2. How did your actual participation level compare to your intentions when you enrolled in this course? I participated …
   - o Less than I originally intended
   - o About the same as I originally intended
   - o More than I originally intended

3. On average, how many hours did you work on this course per week? This includes assignments, course material, and video lectures.

   0 2 4 6 8 10 12 14 16 18 20

4. Did the course meet your expectations? Choose the one that applies the most.
   - o This course exceeded my expectations
   - o This course was exactly what I expected
   - o This course did not meet my expectations

The activities of the instructor(s) and / or teaching assistant(s) form an important part of the online learning experience. How much do you agree with the following statements?

The instructor(s) and / or teaching assistant(s) …

5. … had an active role on the course forums.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

6. … adequately responded to questions from students.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

7. … motivated me to complete this course.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

How much do you agree with the following statements? On the course's discussion forum …

8. … I felt encouraged to ask for help from the course instructor(s)

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

9. … I felt encouraged to discuss the course with other students

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

10. … I felt encouraged to help other students with their course work

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

11. … the answers that my questions received were useful

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

12. … I felt encouraged to ask for help with my course work from other students

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

13. … the questions that I asked were answered

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

The next questions in this survey are about what you learned and took from this course.

14. On a scale from 1 to 10 (1: very poor, 10: excellent), what overall grade would you give this course?

1 2 3 4 5 6 7 8 9 10

15. Which one aspect of this course did you like the most? Please explain.

……

……

16. Which one aspect of this course would you like us to improve? Please explain.

……

……

How much do you agree with the following statements?

17. I would recommend this course to another person.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

18. I would take another course given by this (team of) instructor(s).

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

19. This course helps me advance in my career / professional field.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

20. I learned a lot from this course.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

21. I can apply what I have learned in this course in my daily work.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

22. This course was well worth my time.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

23. What do you consider the most valuable thing you have learned in this course?

    ……

……

24. After participating in this course, what would you like to learn next, related to the subject of this course?

    ……

    ……

We conclude the survey with some questions about the course content: its difficulty, the learning activities during the course, and the quality of the course materials and supporting website.

25. How would you rate the difficulty level of the course?
    - Far too difficult
    - Too difficult
    - About right
    - Too easy
    - Far too easy

26. How would you describe the amount of work required for the course?
    - Far too little
    - Too little
    - About right
    - Too much
    - Far too much

27. How would you describe the duration (in weeks) of the course?
    - Far too short
    - Too short
    - About right
    - Too long
    - Far too long

How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments …

28. … allowed me to identify what I know and can do

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

29. … were too easy

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

30. … reflected the content of the lectures

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

31. … were always clear

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

32. … were too difficult

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

33. … made me apply what I've learned in the course

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

34. Which aspects of the course videos did you especially like and which would you like us to change or improve? Please drag and drop all listed aspects into the appropriate group. You do not need to use all items.

Items	Aspects I liked	Aspects I would like you to improve
Content matter	o	o
Usefulness	o	o
Audio quality	o	o
Understandability	o	o
Subtitles	o	o
Slides readability	o	o
Video quality	o	o
Length	o	o
Animation quality	o	o

35. Which aspects of the course handouts did you especially like and which would you like us to change or improve? Please drag and drop all listed aspects into the appropriate group. You do not need to use all items.

Items	Aspects I liked	Aspects I would like you to improve
Content matter	o	o
Understandability	o	o
Image quality	o	o
Usefulness	o	o
Length	o	o
Readability	o	o
Clarity	o	o

We conclude the survey with some statements about your overall course experience.

How much do you agree with the following statements?

36. The course provided me with the newest insights in research in the field.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

37. The course team created a positive atmosphere to learn.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

38. The course had a right balance between self-study activities and pair/group work.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

39. I was able to access the course from any device (computer, tablet, smartphone, etc.)

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

40. Course team sent updates regularly enough.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

41. The course contained enough exercises to practice with instant feedback.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

42. The course schedule and deadlines allowed me to manage my own time during the course.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

43. Assignments were based on real world scenarios.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

44. If you have any other specific suggestions, praise, or other comments for the course organizers, please feel free to share them below:

……

……

We thank you for your time spent taking this survey.

Appendix 6: Additional questionnaires

Take note, the number of questions to be added to each cohort questionnaire differs. The game-content-group has additional specific questions on the games.

Learner in Default cohort, additional questions

How much do you agree with the following statement? I remained active in the course because…

1. I wanted to earn a certificate

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

2. I wanted to improve my career

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

3. It relates to my academic program

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

4. I had fun

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

5. I found the course interesting

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

6. The course stimulated my curiosity

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

7. I felt challenged

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

8. I felt connected with other students

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

9. I felt connected with the teachers

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

10. I received enough feedback during the course

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

11. There were enough exercises to test my knowledge

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

12. There was enough course material available

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

13. The content's applicability to real-world situations is high

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

Learner in Game Content group, additional questions

How much do you agree with the following statement? I remained active in the course because…

How much do you agree with the following statements? The games…

1. Motivated me to continue with the course

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

2. Were useful

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

3. Were fun

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

4. Were challenging

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

How much do you agree with the following statements?

5. The games reminded me why I enrolled for this course

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

6. I was curious what the next game would be

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

7. I tried to get a higher rank in the games

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

8. I studied extra for the practice multiple choice questions, because I wanted a better multiplier

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

9. The games helped me to understand the course material

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
O	O	O	O	O

If you have any improvements, praise, or other comments regarding the games, please feel free to share them below:

…….

…….

Appendix 7: Data coding schemes

This appendix shows the data coding scheme's used to transform the questionnaire numbers into data that can analysed in SPSS.

Contents

DelftX pre-questionnaire data coding scheme

Source	Survey question number	Description	Variable	Code
DELFTX pre-survey	Q2.3	How did you discover this course? I found out about this course through: - Selected Choice	tpeq.disc	1 = edX website 2 = edX newsletter 3 = TU Delft Homepage 4 = TU Delft Online Learning webpage 5 = Topic search Google 6 = Someone told me 7 = Online article I read 8 = Online advertisement 9 = Social media 10 = Newspaper or magazine 11 = Flyer or poster 12 = Some other way
DELFTX pre-survey	Q2.3_12_TEXT	How did you discover this course? I found out about this course through: - Text		text
DELFTX pre-survey	Q2.4_5_TEXT	Who told you about this course? - Text		text
DELFTX pre-survey	Q2.4	Who told you about this course? - Selected Choice		1 = Fellow student 2 = teacher / professor 3 = Colleague / professional acquaintance 4 = Family or friends 5 = Other
DELFTX pre-survey	Q2.5_6_TEXT	Why did you enroll in this course? Choose the answer that best describes your motivation: - Text		text

DELFTX pre-survey	Q2.5	Why did you enroll in this course? Choose the answer that best describes your motivation: - Selected Choice	tpeq.enrolmot	1 = I can use what I will learn in this course in my daily work activities 2 = I am taking this course to improve my future career prospects 3= I want to get a degree in this field (or related area) 4 = I know the instructors of this course and am very interested in their work 5 = I am interested in this subject and want to learn more about it 6 = Other
DELFTX pre-survey	Q2.6_5_TEXT	How are you pursuing a degree in this field? Choose the answer that best describes your situation: - Text		text
DELFTX pre-survey	Q2.6	How are you pursuing a degree in this field? Choose the answer that best describes your situation: - Selected Choice		1 = I am currently studying this (or a similar) topic at Delft University of Technology. 2 = I am interested in studying this topic at another university. 3 = Other
DELFTX pre-survey	Q2.7_Q3.6_1	How important were the following factors in your decision to enroll in this course? - My general experience with online university courses.	tpeq.imp1	-2 = Unimportant -1 = Of little importance 0 = Moderately important 1 = Important 2 = Very important
DELFTX pre-survey	Q2.7_Q3.6_2	How important were the following factors in your decision to enroll in this course? - My experience with other DelftX courses on edX.	tpeq.imp2	-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_3	How important were the following factors in your decision to enroll in this course? - My experience with other online Delft University of Technology / DelftX courses.		-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_4	How important were the following factors in your decision to enroll in this course? - The status and reputation of Delft University of Technology.	tpeq.imp3	-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_5	How important were the following factors in your decision to enroll in this course? - The	tpeq.imp4	-2 = Unimportant …. 1 = Important

		possibility of earning a Verified Certificate.		
DELFTX pre-survey	Q2.7_Q3.6_6	How important were the following factors in your decision to enroll in this course? - The possibility of earning European Credits and a certificate by completing this course.		-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_7	How important were the following factors in your decision to enroll in this course? - The possibility of earning Continuous Education Units(CEUs) and a certificate by completing this course.		-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_8	How important were the following factors in your decision to enroll in this course? - The relevance of the course topic for my occupation or work.	tpeq.imp5	-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.7_Q3.6_9	How important were the following factors in your decision to enroll in this course? - A recommendation from someone else.	tpeq.imp6	-2 = Unimportant …. 1 = Important
DELFTX pre-survey	Q2.8	How familiar were you with Delft University of Technology prior to this course? Choose the answer that best describes your familiarity with Delft University of Technology.	tpeq.familiarTU	1 = Not at all familiar 2 = Slightly familiar 3 = Somewhat familiar 4 = Moderately familiar 5 = Very Familiar
DELFTX pre-survey	Q3.3	How many online courses did you complete?	tpeq.prvcourscompl	0 = 0 … 5 = 5 or more
DELFTX pre-survey	Q3.2	How many online courses have you taken before?	tpeq.prvcoursdone	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5 = 5 or more
DELFTX pre-survey	Q3.4_1	On average, how many hours per week can you dedicate to this course? – Hours	tpeq.hourdedic.	ratio

DELFTX pre-survey	Q3.5_4_TEXT	How do you intend to participate in this course? Choose the answer that most closely describes your intentions: - Text		text
DELFTX pre-survey	Q3.5	How do you intend to participate in this course? Choose the answer that most closely describes your intentions: - Selected Choice	tpeq.intention	1 = I want to complete this course 2 = I will only do the specific parts of this course that interest me 3 = I am just checking this course out 4 = Other
DELFTX pre-survey	Q3.6_1	Which activities do you plan on doing in this course? Will you ... - ... watch the videos?	tpeq.activplans1	-1 = No 0 = Don't know yet 1 = Yes
DELFTX pre-survey	Q3.6_2	Which activities do you plan on doing in this course? Will you ... - ... do assignments and quizzes?	tpeq.activplans2	-1 = No 0 = Don't know yet 1 = Yes
DELFTX pre-survey	Q3.6_3	Which activities do you plan on doing in this course? Will you ... - ... participate in the forums?	tpeq.activplans3	-1 = No 0 = Don't know yet 1 = Yes
DELFTX pre-survey	Q3.6_4	Which activities do you plan on doing in this course? Will you ... - ... read course materials?	tpeq.activplans4	-1 = No 0 = Don't know yet 1 = Yes
DELFTX pre-survey	Q3.8	How would you rate your level of English proficiency?	tpeq.engprof	-2 = Poor -1 = Fair 0 = Average 1 = Good 2 = Very good
DELFTX pre-survey	Q3.9	How comfortable are you communicating in English?	tpeq.engcomfort	-2 = Very uncomfortable -1 = Uncomfortable 0 = Neither comfortable nor uncomfortable 1 = Comfortable 2 = Very comfortable
DELFTX pre-survey	Q5.2_1	What is your age? - Years	tpeq.age	ratio
DELFTX pre-survey	Q5.3	What is your gender?	tpeq.gender	1 = Male 2 = Female 3 = Other
DELFTX pre-survey	Q5.4_1	What is your (first) nationality? Please select the corresponding country. – Continent	tpeq.continent	6 choices
DELFTX pre-survey	Q5.4_2	What is your (first) nationality? Please select the corresponding country. – Country	tpeq.country	63 choices
DELFTX pre-survey	Q5.7_9_TEXT	What is the highest degree or level of education you have completed? – Text		Text

DELFTX pre-survey	Q5.7	What is the highest degree or level of education you have completed? - Selected Choice	tpeq.degree	1 = Elementary school degree 2 = Secondary school 3 = Collge / Associate's degree 4 = Professional / vocational degree 5 = Bachelor's degree 6 = Master's degree 7 = Postgraduate / graduate degree 8 = Ph.D.
DELFTX pre-survey	Q5.8_7_TEXT	Which of the following best describes your current situation? – Text		Text
DELFTX pre-survey	Q5.8	Which of the following best describes your current situation? - Selected Choice	tpeq.occupt.	1 = Student 2 = Recently graduated 3 = Working professional 4 = Lookign for a job 5 = Parent / care-giver 6 = Retired 7 = Other
DELFTX pre-survey	Q5.9	Which of the following most closely matches your job title? - Selected Choice	tpeq.jobtitle	1 = Intern 2 = Entry level / trainee 3 = Analyst / Associate 4 = Developer / programmer 5 = Engineer / Technician 6 = Lecturer 7 = Professor (incl. associate, assistant) 8 = Consultant 9 = Researcher 10 = Project / Team leader 11 = Manger / Director 12 = Executive / President 13 = Business Owner / Entrepreneur 14 = Freelancer / Self-employed 15 = Other
DELFTX pre-survey	Q5.9_15_TEXT	Which of the following most closely matches your job title? – Text		Text
DELFTX pre-survey	Q5.10_15_TEXT	Which of the following most closely matches your most recent job title? - Text		Text
DELFTX pre-survey	Q5.10	Which of the following most closely matches your most recent job title? - Selected Choice	tpeq.recentjobtitle	1 = Intern ... 15 = Other
DELFTX pre-survey	Q5.11_1	In which industry do you currently work? – Sector	tpeq.sector	63 choices
DELFTX pre-survey	Q5.11_2	In which industry do you currently work? – Industry	tpeq.industry	63 choices

DELFTX pre-survey	Q5.12_1	On average, how many hours do you work per week? – Hours	tpeq.workhourscareer	ratio
DELFTX pre-survey	Q5.13	Are you allowed to follow this course during working hours?	tpeq.allowtofollow	4 = Yes 5 = No 6 = Don't know
DELFTX pre-survey	Q5.14_1	How many years of working experience do you have? – Years	tpeq.workexp	ratio

DelftX mid-questionnaire data coding scheme

Source	Survey question number	Description	Variable	Code
DelftX mid-survey	Q2	Please let us know how much you're enjoying the course through the smiley face (click on the smiley to activate it).	tmq.enjoy	1 tot 5
DelftX mid-survey	Q4	Which aspect of this course do you especially like? Please name one.		text
DelftX mid-survey	Q3	Which aspect of this course would you like us to change or improve? Please name one.		text

DelftX Post-questionnaire data coding scheme

Source	Survey question number	Description	Variable	Code
DelftX post-questionnaire	Q1.1	If you do not wish to participate in this evaluation, please decline participation by selecting the "Disagree" button.		
DelftX post-questionnaire	Q2.3	Since the start of the course, how would you describe your participation level? In order for the survey to unfold correctly you need to answer this question, but don't worry, all the other questions are optional!	tpoq.parti	1 = I did not participate in this course at all 2 = I did not participate in this course, but i did browse around a little bit 3 = I only looked at specific parts of the course that I was interested in 4 = I participated in the coruse, but I stopped participating along the way 5 = I participated in the course up until the end
DelftX post-questionnaire	Q2A.2	Could you please describe which specific parts of the course you were interested in and why?		Text
DelftX post-questionnaire	Q2A.1	Could you please describe the reason(s) why you did not start the course?		Text
DelftX post-questionnaire	Q2B.1_6_TEXT	Could you please describe the reason(s) why you decided not to continue with the course? Choose the one that applies the most. - Other â€¦ - Text		Text
DelftX post-questionnaire	Q2B.1	Could you please describe the reason(s) why you decided not to continue with the course? Choose the one that applies the most. - Selected Choice		1 = Personal obligations 2 = Due to an unexpected change in available time 3 = Due to a general lack of time 4 = The course was not what I expected 5 = The course was too difficult for me 6 = Other
DelftX post-questionnaire	Q2B.2	How did your actual participation level compare to your intentions when you enrolled in this course? I participated â€¦	tpoq.parti.comp	1 = Less than I originally intended 2 = About the same as I originally intended 3 = More than I originally intended
DelftX post-questionnaire	Q2B.4	Could you please describe the reason(s) why you participated more than you originally intended?		Text
DelftX post-questionnaire	Q2B.3	Could you please describe the reason(s) why you participated less than you originally intended?		Text
DelftX post-questionnaire	Q3.1_1	On average, how many hours did you work on this course per week? This includes assignments, course material, and video lectures. - Hours:	tpoq.hours.week	ratio

DelftX post-questionnaire	Q3.2	Did the course meet your expectations? Choose the one that applies the most.	tpoq.expect	1 = This coursre did not meet my expectations 2 = This course was exactly what I expected 3 = This course exceeded my expectations
DelftX post-questionnaire	Q3.4	Could you please describe how the course did not meet your expectations?		Text
DelftX post-questionnaire	Q3.3	Could you please describe how the course exceeded your expectations?		Text
DelftX post-questionnaire	Q3.5_1	The activities of the instructor and / or teaching assistants form an important part of the online learning experience. How much do you agree with the following statements? The instructor and / or teaching assistants had an active role on the course forums.	tpoq.int.act.forum	-2 = Strongly disagree -1 = Disagree 0 = Neither agree nor disagree 1 = Agree 2 = Strongly agree
DelftX post-questionnaire	Q3.5_2	The instructor and / or teaching assistants adequately responded to questions from students.	tpoq.int.respond .forum	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.5_3	The instructor and / or teaching assistants motivated me to complete this course.	tpoq.int.mot	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_6	How much do you agree with the following statements? On the course's discussion forum I felt encouraged to discuss the course with other students	tpoq.int.dicuss	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_7	I felt encouraged to help other students with their course work	tpoq.int.tohelp	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_8	I felt encouraged to ask for help with my course work from other students	tpoq.int.askhelp	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_9	I felt encouraged to ask for help from the course instructors	tpoq.int.askhelp	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_10	the questions that I asked were answered	tpoq.int. questanswer	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q3.6_11	the answers that my questions received were useful	tpoq.int. answer.usef	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.2_1	On a scale from 1 to 10 (1: very poor, 10: excellent), what overall grade would you give this course? - Score:	tpoq. gradeforcourse	1 …10
DelftX post-questionnaire	Q4.4	Which one aspect of this course would you like us to improve? Please explain.		Text
DelftX post-questionnaire	Q4.3	Which one aspect of this course did you like the most? Please explain.		Text
DelftX post-questionnaire	Q4.5_Q5.5_1	How much do you agree with the following statements? - I learned a lot from this course.	tpoq.pmq. learnlot	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_2	How much do you agree with the following statements? - I can apply what I have learned in this course in my daily work.	tpoq.em.appl	-2 = Strongly disagree … 2 = Strongly agree

DelftX post-questionnaire	Q4.5_Q5.5_3	How much do you agree with the following statements? - This course helps me advance in my career / professional field.	tpoq.em.careeradv	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_4	How much do you agree with the following statements? - This course was well worth my time.	tpoq.pmq. worthmytime	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_5	How much do you agree with the following statements? - I would recommend this course to another person.	tpoq.pmq.reccom	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_6	How much do you agree with the following statements? - I would take another course given by this (team of) instructor(s).	tpoq.int. othercourse	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_7	How much do you agree with the following statements? - I would take another course from DelftX.	tpoq.othercourse. DelftX	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.5_Q5.5_8	How much do you agree with the following statements? - I would consider applying for a professional course by DelftX in a related area.	tpoq.proff.DelftX	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q4.7	What do you consider the most valuable thing you have learned in this course?		Text
DelftX post-questionnaire	Q5.3	Could you please describe why you think the course was too difficult?		Text
DelftX post-questionnaire	Q5.2	How would you rate the difficulty level of the course?	tpoq.pmq.diff	-2 = Far too difficult -1 = Too difficult 0 = About right 1 = Too easy 2 = Far too easy
DelftX post-questionnaire	Q5.5	How would you describe the amount of work required for the course?	tpoq.pmq.workreq	-2 = Far too little -1 = Too little 0 = About right 1 = Too much 2 = Far too much
DelftX post-questionnaire	Q5.4	Could you please describe why you think the course was too easy?		Text
DelftX post-questionnaire	Q5.6	How would you describe the duration (in weeks) of the course?	tpoq.pmq.durat	-2 = Far too short -1 = Too short 0 = About right 1 = Too long 2 = Far too long
DelftX post-questionnaire	Q5.7_11	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments were always clear	tpoq.pmq.assclear	-2 = Strongly disagree -1 = Disagree 0 = Neither agree nor disagree 1 = Agree 2 = Strongly agree
DelftX post-questionnaire	Q5.7_12	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments were too easy	tpoq.pmq.asseasy	-2 = Strongly disagree … 2 = Strongly agree

DelftX post-questionnaire	Q5.7_13	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments were too difficult	tpoq.pmq.assdiff	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q5.7_14	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments reflected the content of the lectures	tpoq.pmq. assreflcontent	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q5.7_15	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments made me apply what I've learned in the course	tpoq.pmq. assmadeapply	-2 = Strongly disagree … 2 = Strongly agree
DelftX post-questionnaire	Q5.7_16	How much do you agree with the following statements about assignments in the course (quizzes, homework, peer review assignments etc.)? The assignments allowed me to identify what I know and can do	tpoq.pmq. assmade.iden	-2 = Strongly disagree … 2 = Strongly agree

Additional questionnaire data coding scheme

Only the data coding scheme for the game group is presented, as the first 13 questions are the same for the control and game group, while the game group has extra game evaluation questions.

Source	Survey question number	Description	Variable	Code
Add. Game-survey	Q1_1	I wanted to earn a certificate	aq.em.certf	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strong agree = 5
Add. Game-survey	Q1_2	I wanted to improve my career	aq.em.career	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_3	It relates to my academic program	aq.em.academic	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_4	I had fun	aq.im.fun	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_5	I found the course interesting	aq.im.interest	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_6	The course stimulated my curiosity	aq.im.curious	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_7	I felt challenged	aq.im.challenge	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_8	I felt connected with other students	aq.int.stud	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_9	I felt connected with the teachers	aq.int.teach	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_10	I received enough feedback during the course	aq.int.feedback	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_11	There were enough exercises to test my knowledge	aq.pmq.exercise	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_12	There was enough course material available	aq.pmq.material	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q1_13	The content's applicability to real-world situations is high	aq.pmq.applicable	Strongly disagree = 1 … Strongly agree = 5
Add. Game-survey	Q3_1	The games motivated me to continue with the course	aq.g.motiv	Strongly disagree = 1 … Strongly agree = 6
Add. Game-survey	Q3_2	The games were useful	aq.g.useful	Strongly disagree = 1 … Strongly agree = 7
Add. Game-survey	Q3_3	The games were fun	aq.g.fun	Strongly disagree = 1 … Strongly agree = 8
Add. Game-survey	Q3_4	The games were challenging	aq.g.challenging	Strongly disagree = 1 … Strongly agree = 9
Add. Game-survey	Q4_1	The games reminded me why I enrolled for this course	aq.g.enroll	Strongly disagree = 1 … Strongly agree = 10
Add. Game-survey	Q4_2	I was curious what the next game would be	aq.g.curious	Strongly disagree = 1 … Strongly agree = 11

Add. Game-survey	Q4_3	I tried to get a higher rank in the games	aq.g.highrank	Strongly disagree = 1 … Strongly agree = 12
Add. Game-survey	Q4_4	I studied extra for the practice multiple-choice questions, because I wanted a better multiplier	aq.g.multpl	Strongly disagree = 1 … Strongly agree = 13
Add. Game-survey	Q4_5	The games helped me to understand the course material	aq.g.underst	Strongly disagree = 1 … Strongly agree = 14
Add. Game-survey	Q5	If you have any improvements, praise, or other comments regarding the games please feel free to share them below		text

Grade report data coding scheme

Source	Survey question number	Description	Variable	Code
edX grade report	NA	Grade	Grade	ratio
edX grade report	NA	Average grade Homework	HomeworkAVG	ratio
edX grade report	NA	Number of Homework Made	HomeworkMade	ratio
edX grade report	NA	Multiple choice 1: Homework	Homework1	ratio
edX grade report	NA	Multiple choice 2: Homework	Homework2	ratio
edX grade report	NA	Multiple choice 3: Homework	Homework3	ratio
edX grade report	NA	Multiple choice 4: Homework	Homework4	ratio
edX grade report	NA	Multiple choice 5: Homework	Homework5	ratio
edX grade report	NA	Multiple choice (Avg)	MC.avg	ratio
edX grade report	NA	Cohort Name	Cohort	Control Cohort = 1 Game Cohort = 2 Default Cohort = 3 MOOC for Credits = 4 Blank = 5
edX grade report	NA	Enrollment Track	enroll.track	Audit = 1 Verified = 2
edX grade report	NA	Verification Status	Verf.status	N/A = 0 ID verified = 1 Not ID verified = 2
edX grade report	NA	Certificate Eligible	Cert.eligible	N = 1 Y = 2
edX grade report	NA	Certificate Delivered	Cert.deliverd	N = 1 Y = 2
edX grade report	NA	Certificate Type	Cert.type	N/A = 0 Verified = 1
edX grade report	NA	Enrollment Status	Enroll.status	enrolled = 1 unenrolled = 2

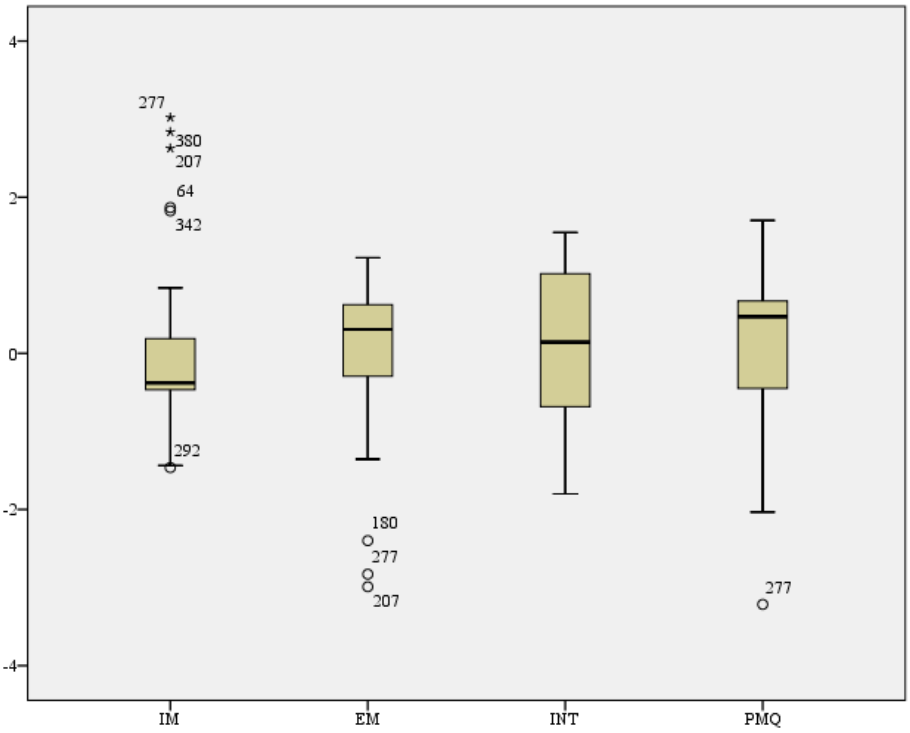
Game data grading coding scheme

Source	Survey question number	Description	Variable	Code
InThere	NA	Number of games played in total	GamesPlayed	ratio
InThere	NA	Times Game 1 is played	TimesPlayedG1	ratio
InThere	NA	Times Game 2 is played	TimesPlayedG2	ratio
InThere	NA	Times Game 3 is played	TimesPlayedG3	ratio
InThere	NA	Times Game 4 is played	TimesPlayedG4	ratio
InThere	NA	Times Game 5 is played	TimesPlayedG5	ratio
InThere	NA	Highscore Game 1	HighscoreG1	ratio
InThere	NA	Highscore Game 2	HighscoreG2	ratio
InThere	NA	Highscore Game 3	HighscoreG3	ratio
InThere	NA	Highscore Game 4	HighscoreG4	ratio
InThere	NA	Highscore Game 5	HighscoreG5	ratio

Appendix 8: Boxplots, outlier check

In this appendix the boxplots are shown for the four factors to identify outliers. Observations are called an outlier if it falls more than 1.5(IQR) above the upper quartile or more than 1.5(IQR) below the lower quartile (Agresti & Finlay, 2009). The upper and lower quartiles are top and bottom of the box. It the boxplot the following outliers are identified: 64, 180, 207, 277, 342, and 380.

Looking at the respondents more closely, it was found that respondent 277 filled in only 1's in the additional questionnaire for every question, explaining why this observations is visible as an outlier in three boxplots. The other respondents did vary in their responses, and just appear to rate the MOOC experience more negative. There were not coding mistakes found for any of the respondents. Furthermore, when looking at the grades it was found that the respondents scored a 78%, 50%, 95%, 30%, 15%, and 68% respectively. These grade make it unlikely the respondents rated everything negative out of frustration with their grade. To be sure the outliers did not have any effect on the results of the study, the analyses were conducted without the outliers as well. This did not lead to a different result, showing that the results were not sensitive to the outliers. These previous was seen as enough reason to not exclude the outliers from the dataset.



Appendix 9: Factor analysis

This appendix shows the correlation matrix, KMO and Bartlett's Test and Factor correlation matrix.

The KMO value should be above 0.6 and The Bartlett's Test of sphericity tests if all diagonal elements are 1 and the off diagonal limits are 0 ("Factor Analysis | SPSS Annotated Output," 2017), which is not what you want as this means no correlations exists between the variables in the correlation matrix (next page). The $p < 0.001$ shows that this is not the case. The factor correlation matrix depicts the correlations between the factors.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.			0.709
Bartlett's Test of Sphericity	Approx. Chi-Square		312.771
	df		55
	Sig.		0.000

Factor Correlation Matrix

Factor	PMQ	INT	EM	IM
PMQ	1.000	0.239	0.348	-0.486
INT	0.239	1.000	0.237	-0.305
EM	0.348	0.237	1.000	-0.426
IM	-0.486	-0.305	-0.426	1.000

Extraction Method: Principal Axis Factoring.

Rotation Method: Oblimin with Kaiser Normalization.

Correlation Matrix												
		Certificate	Career	Academic	Fun	Interest	Curiosity	Challenge	Teacher	Feedback	Exercises	Material
Correlation	Certificate	1.000	0.393	0.42	0.111	0.292	0.159	0.166	0.117	0.195	0.101	0.215
	Career	0.393	1.000	0.464	0.308	0.486	0.331	0.431	0.187	0.322	0.066	0.313
	Academic	0.42	0.464	1.000	0.302	0.379	0.426	0.17	0.185	0.373	0.243	0.341
	Fun	0.111	0.308	0.302	1.000	0.642	0.642	0.248	0.534	0.549	0.335	0.379
	Interest	0.292	0.486	0.379	0.642	1.000	0.821	0.494	0.213	0.489	0.423	0.567
	Curiosity	0.159	0.331	0.426	0.642	0.821	1.000	0.441	0.08	0.367	0.382	0.513
	Challenge	0.166	0.431	0.17	0.248	0.494	0.441	1.000	0.158	0.293	0.524	0.533
	Teacher	0.117	0.187	0.185	0.534	0.213	0.08	0.158	1.000	0.668	0.215	0.157
	Feedback	0.195	0.322	0.373	0.549	0.489	0.367	0.293	0.668	1.000	0.302	0.412
	Exercises	0.101	0.066	0.243	0.335	0.423	0.382	0.524	0.215	0.302	1.000	0.741
	Material	0.215	0.313	0.341	0.379	0.567	0.513	0.533	0.157	0.412	0.741	1.000

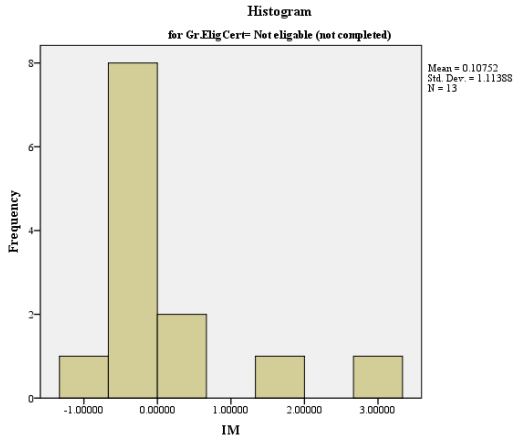
Appendix 10: Part 1 Normality check, factors vs. completion

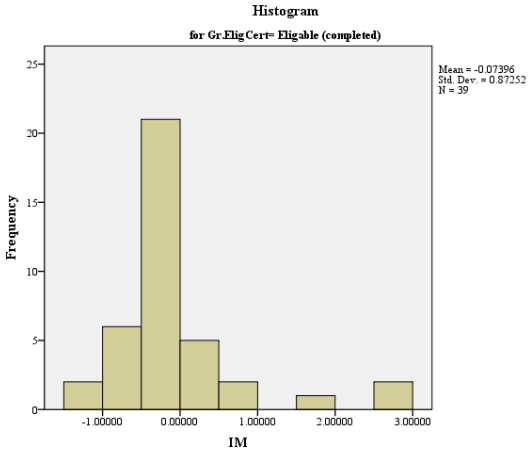
This appendix shows the Skewness, Kurtosis, and Shapiro-Wilk values for every factor in the group of students who completed the MOOC and the group of students who did not. Furthermore, the histogram, Q-Q plots, and Boxplots are shown for the visual inspection of approximate normality.

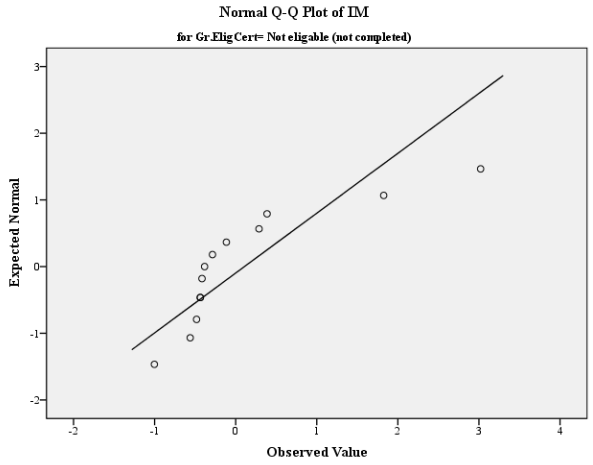
Contents

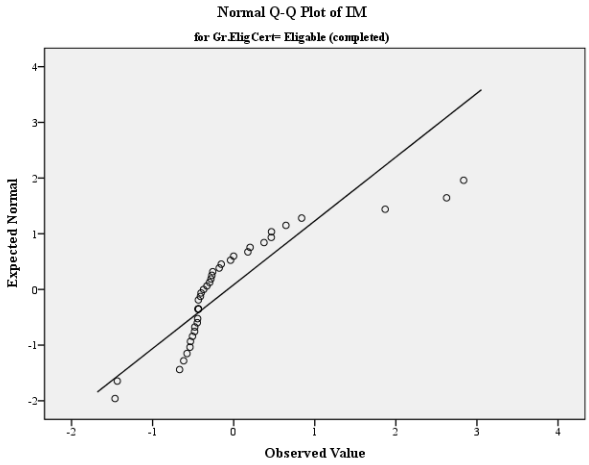
Factor	Grouping variable	Distribution	Statistic	SE	Shapiro-Wilk (p <)
IM	Not completed	Skewness	1.945	0.616	0.002
		Kurtosis	3.499	1.191	
	Completed	Skewness	1.946	0.378	0.000
		Kurtosis	4.621	0.741	
EM	Not completed	Skewness	-1.402	0.616	0.004
		Kurtosis	0.829	1.191	
	Completed	Skewness	-1.978	0.378	0.000
		Kurtosis	5.793	0.741	
INT	Not completed	Skewness	-0.708	0.616	0.010
		Kurtosis	-1.132	1.191	
	Completed	Skewness	-0.378	0.378	0.012
		Kurtosis	-0.907	0.741	
PMQ	Not completed	Skewness	-1.352	0.616	0.001
		Kurtosis	0.643	1.191	
	Completed	Skewness	-0.691	0.378	0.003
		Kurtosis	0.071	0.741	

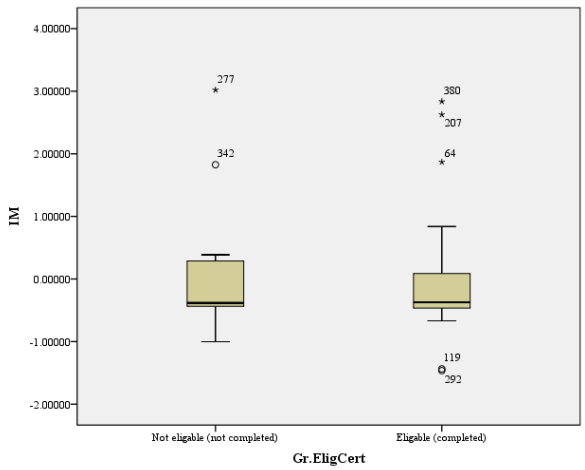
IM vs completion visual normality check



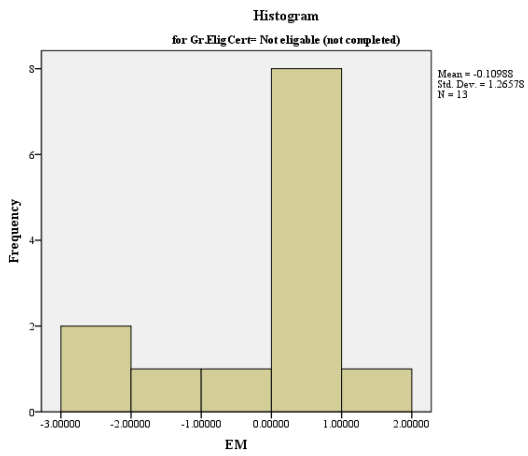


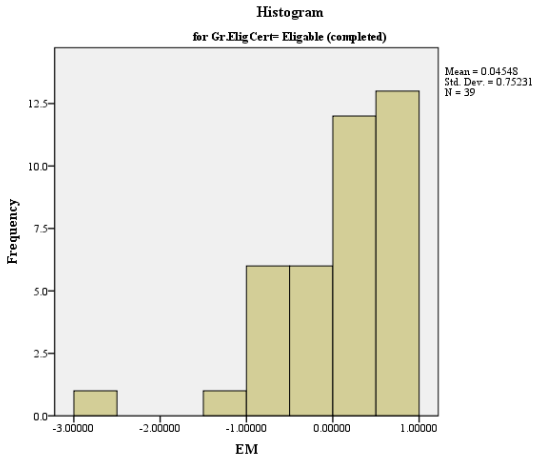


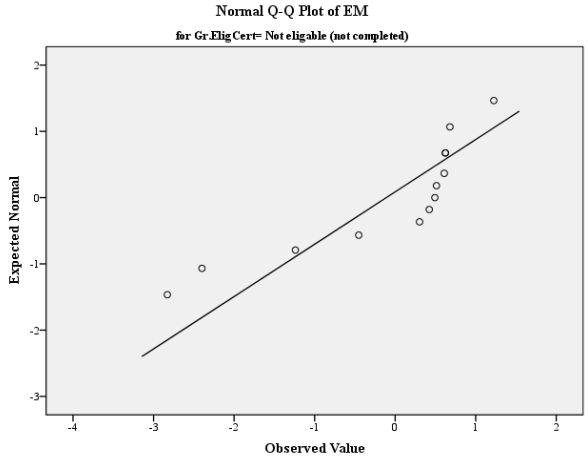


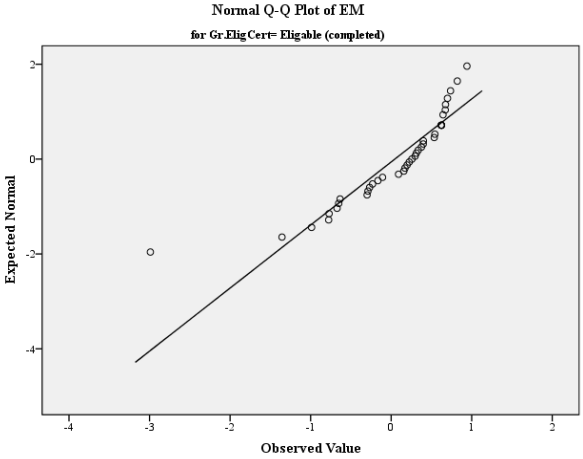


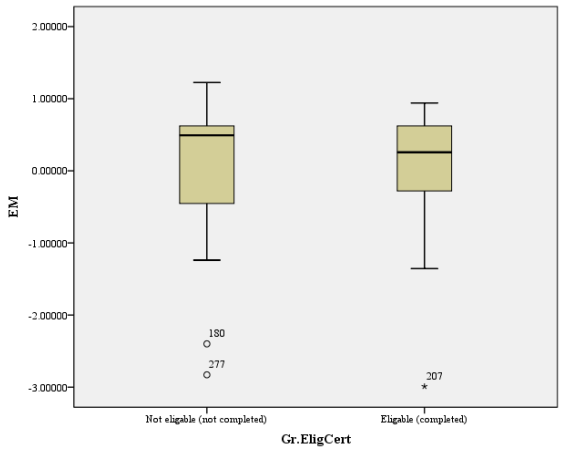
EM vs completion visual normality check



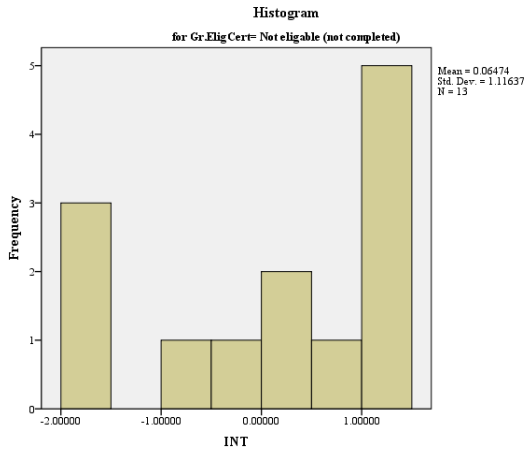


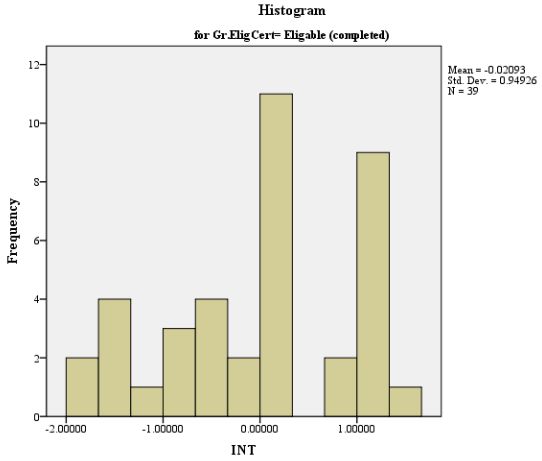


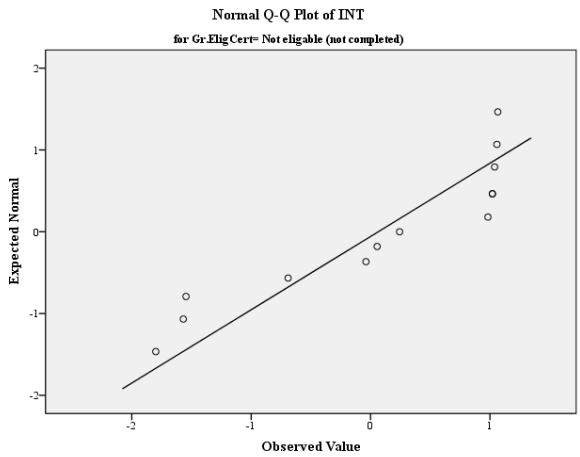


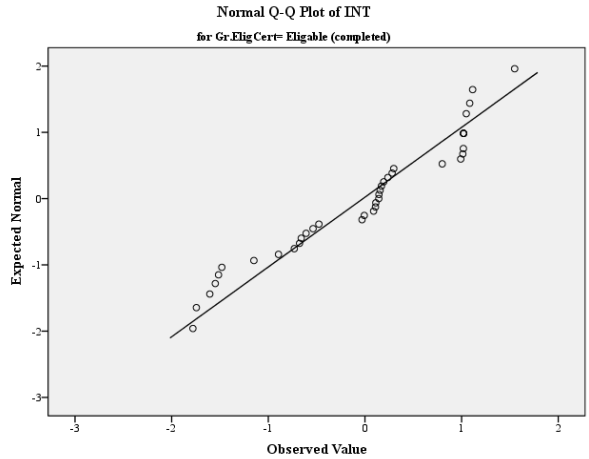


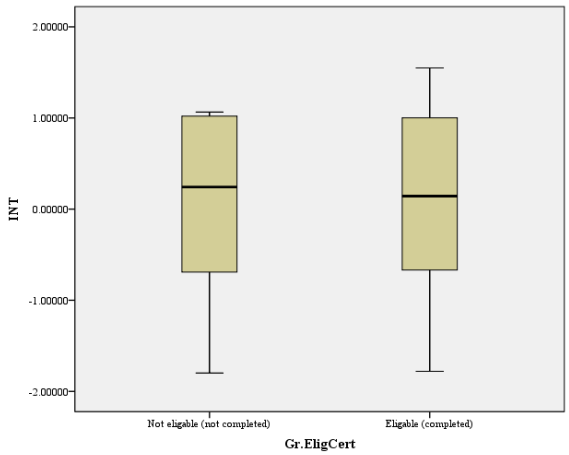
INT vs completion visual normality check



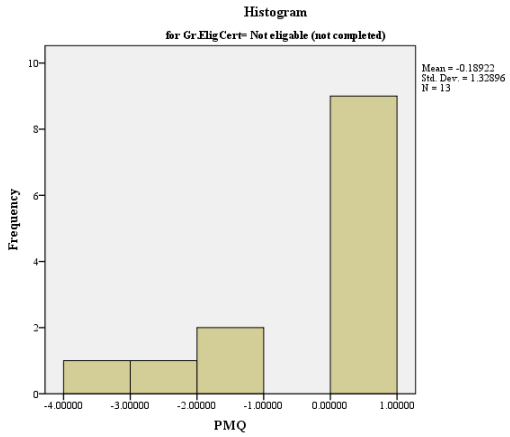


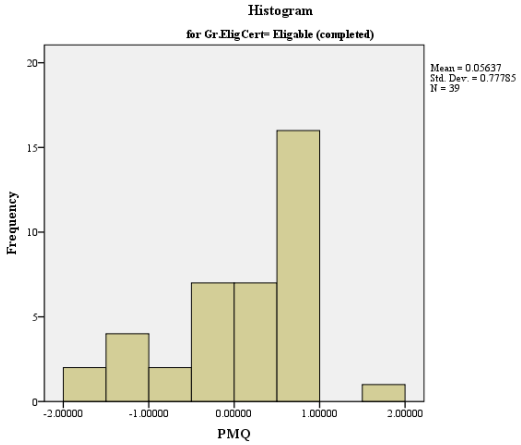


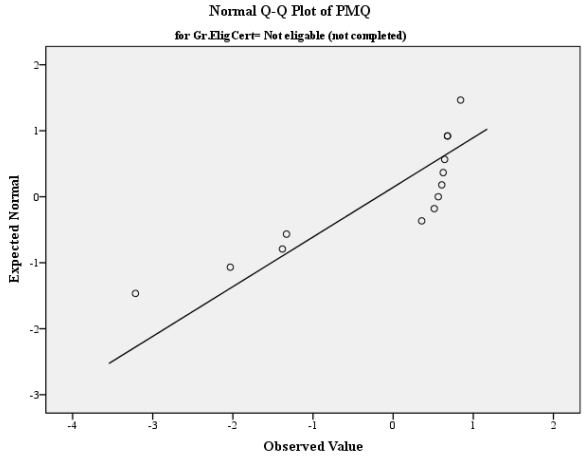


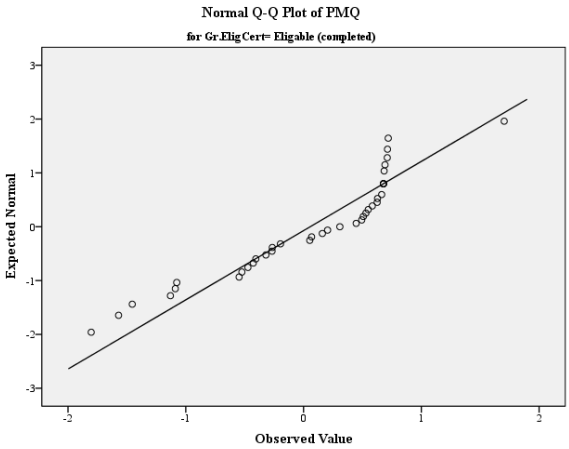


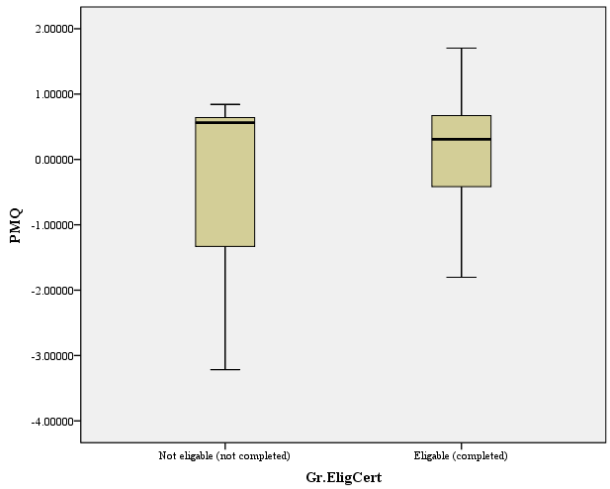
PMQ vs completion normality check



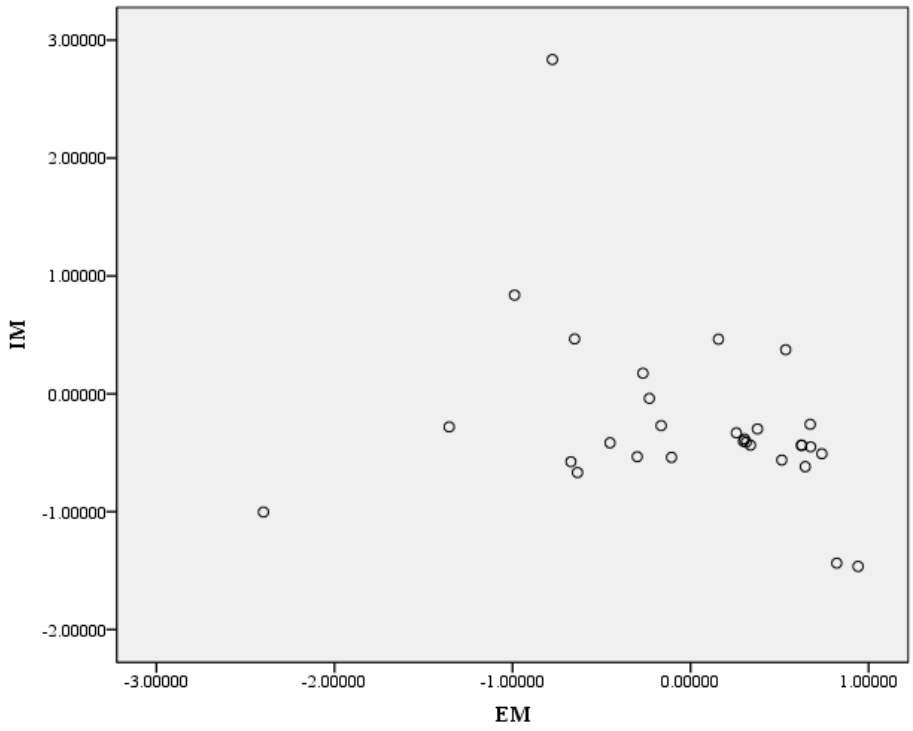


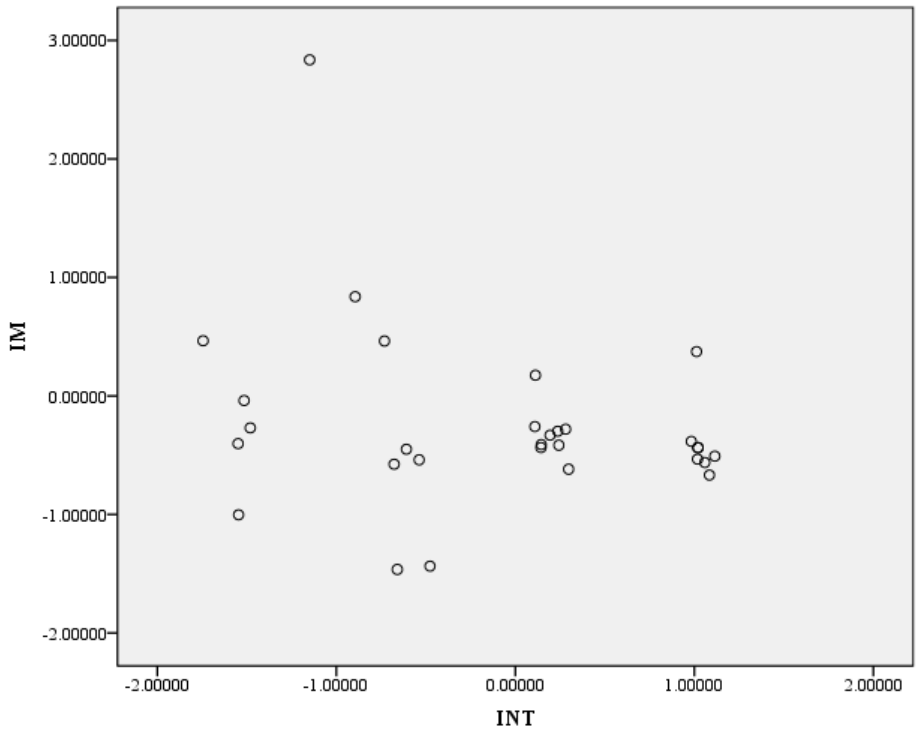


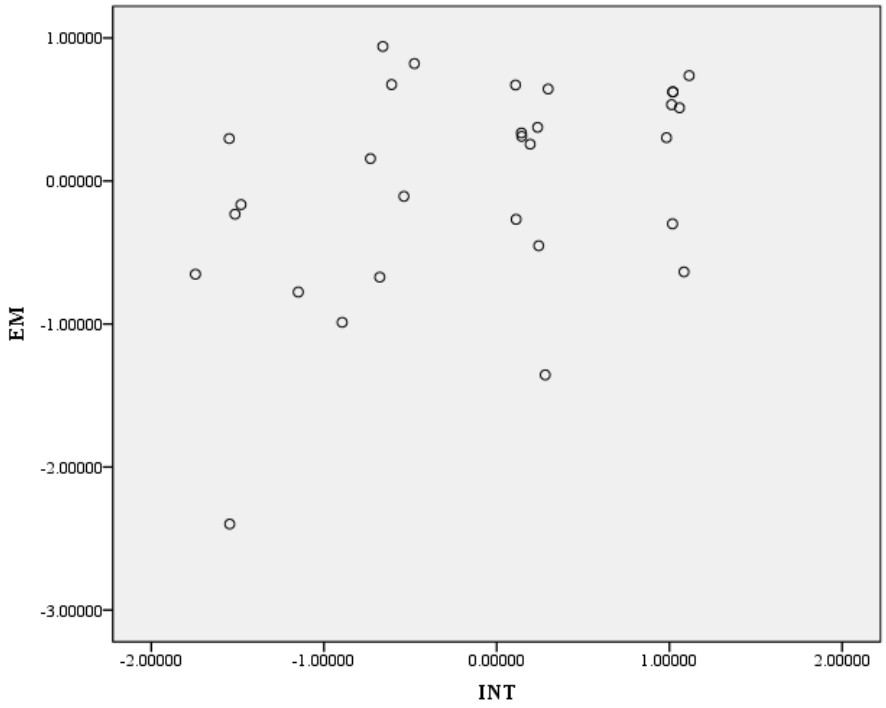


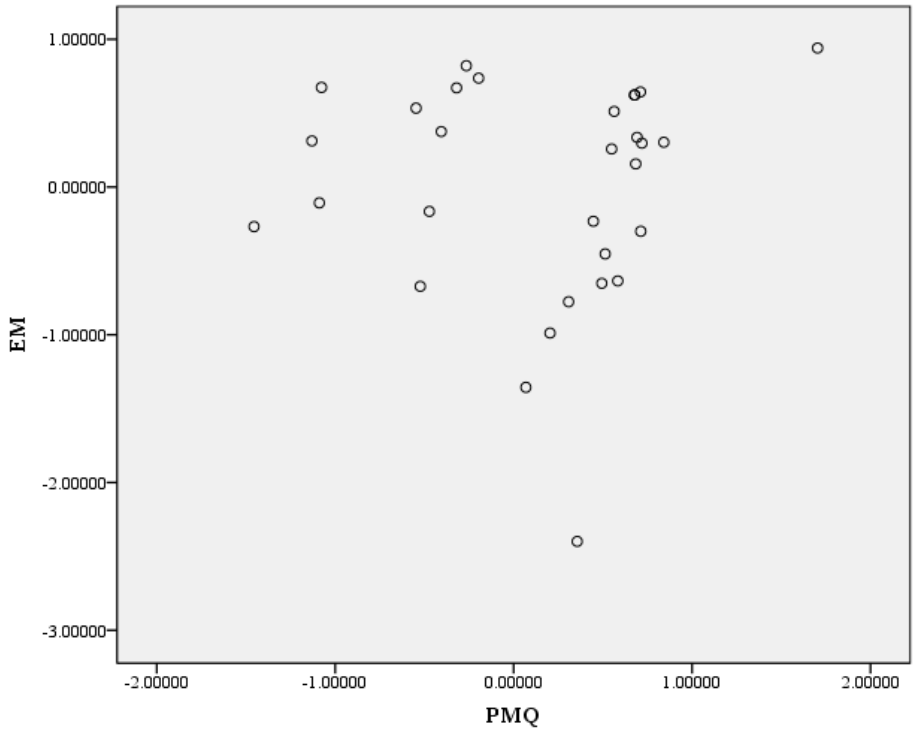


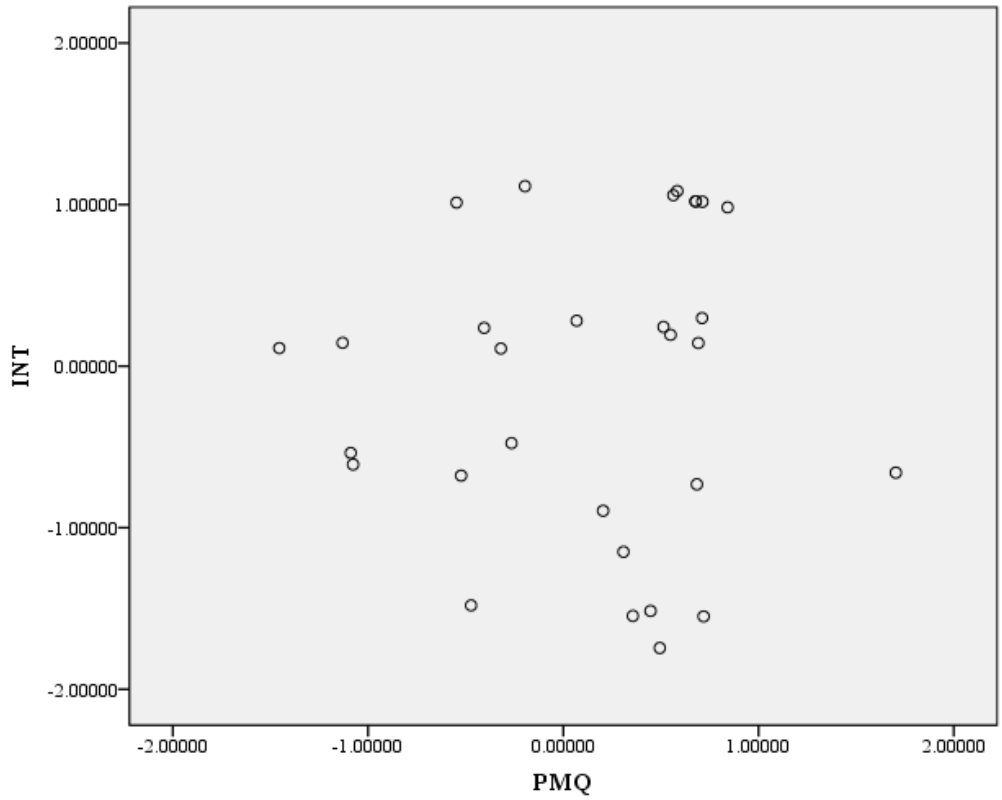
Appendix 11: Part 1 Linearity check for Factors vs Factors






1.00000
0.00000
-1.00000
-2.00000
-3.00000
EM
-2.00000
-1.00000
0.00000
1.00000
2.00000
INT


1.00000
0.00000
-1.00000
-2.00000
-3.00000
EM
-2.00000
-1.00000
0.00000
1.00000
2.00000
PMQ


2.00000
1.00000
0.00000
-1.00000
-2.00000
INT
-2.00000
-1.00000
0.00000
1.00000
2.00000
PMQ

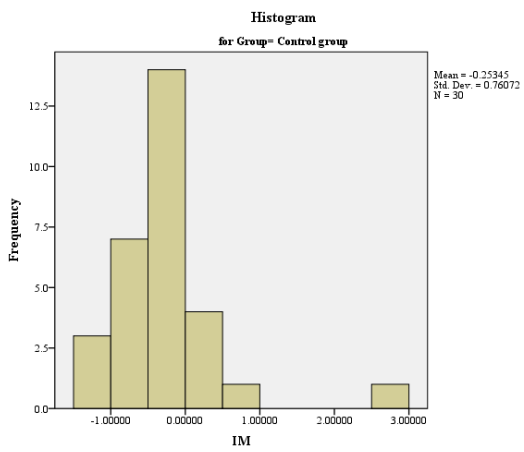
Appendix 12: Part 2 Normality check for SG vs Factors

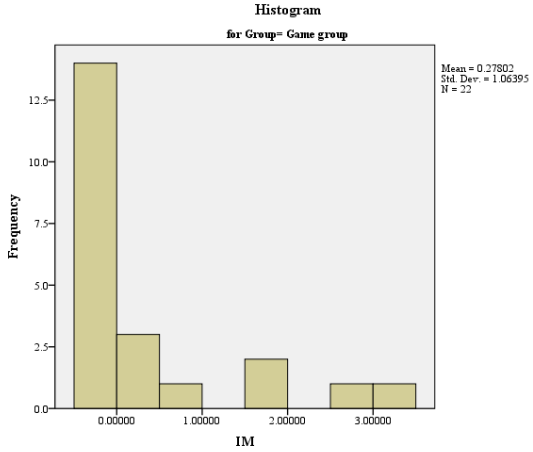
This appendix shows the Skewness, Kurtosis, and Shapiro-Wilk values for every factor in the game group and the control group. Furthermore, the histogram, Q-Q plots, and Boxplots are shown for the visual inspection of approximate normality.

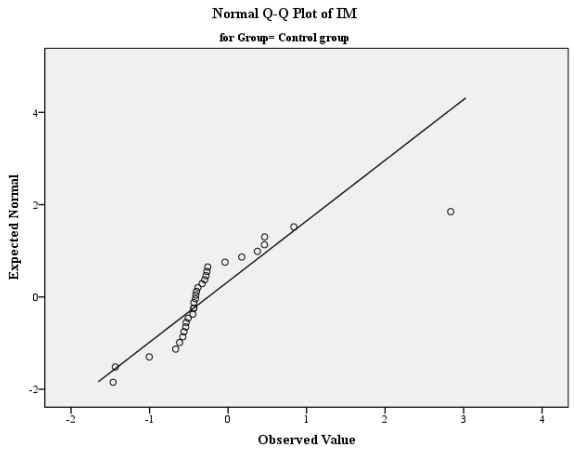
Contents

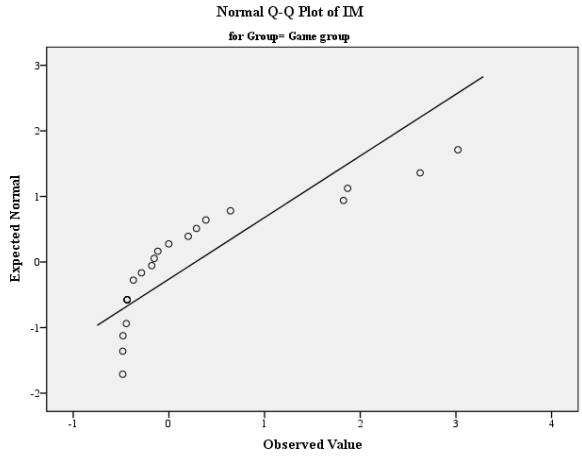
Factor	Grouping variable	Distribution	Statistic	SE	Shapiro-Wilk (p <)
IM	Control group	Skewness	2.328	0.427	0.000
		Kurtosis	9.047	0.833	
	Game group	Skewness	1.630	0.491	0.000
		Kurtosis	1.557	0.953	
EM	Control group	Skewness	-1.300	0.427	0.006
		Kurtosis	2.140	0.833	
	Game group	Skewness	-2.028	0.491	0.000
		Kurtosis	3.634	0.953	
INT	Control group	Skewness	-0.231	0.427	0.015
		Kurtosis	-1.172	0.833	
	Game group	Skewness	-0.825	0.491	0.004
		Kurtosis	-0.473	0.953	
PMQ	Control group	Skewness	-0.478	0.427	0.031
		Kurtosis	-0.112	0.833	
	Game group	Skewness	-1.248	0.491	0.000
		Kurtosis	0.569	0.953	

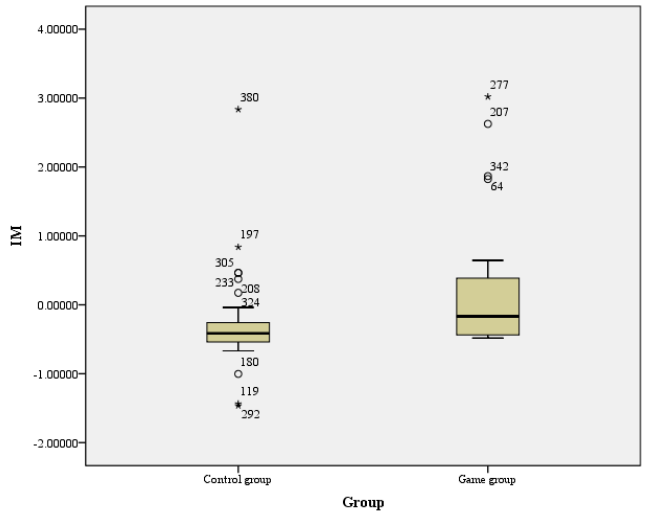
SG vs IM


Histogram
for Group= Control group
Mean = -0.25345
Std. Dev. = 0.76072
N = 30
Frequency
IM

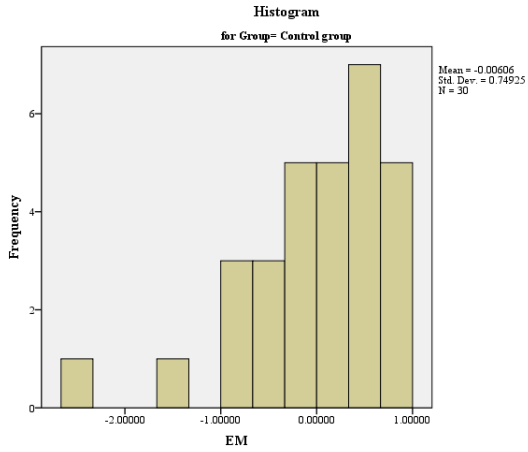

Histogram
for Group= Game group
Mean = 0.27802
Std. Dev. = 1.06395
N = 22
Frequency
IM

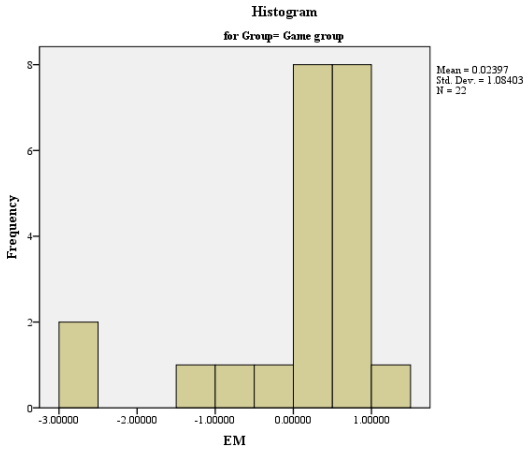

Normal Q-Q Plot of IM
for Group= Control group
Expected Normal
Observed Value

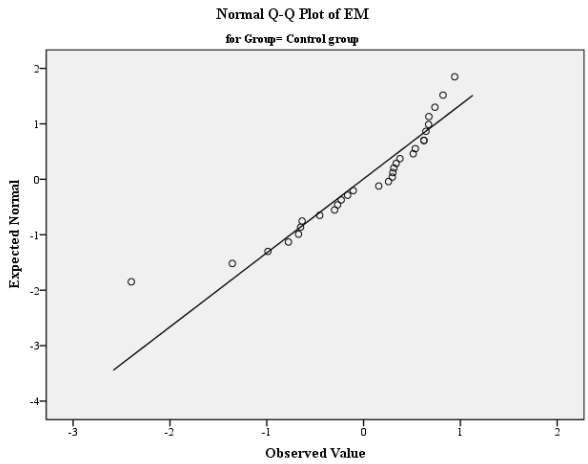

Normal Q-Q Plot of IM
for Group= Game group
Expected Normal
Observed Value

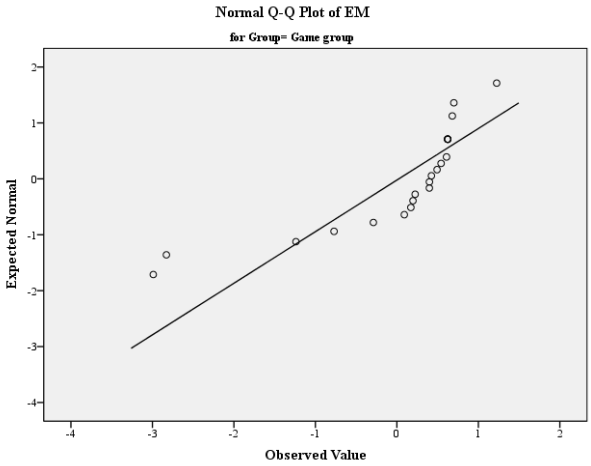

IM
Control group
Game group
Group

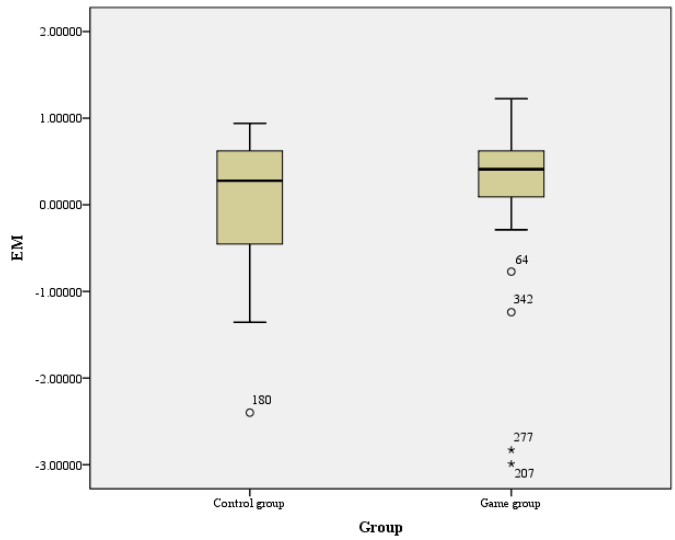
SG vs EM


Histogram
for Group= Control group
Mean = -0.00606
Std. Dev. = 0.74925
N = 30
Frequency
EM

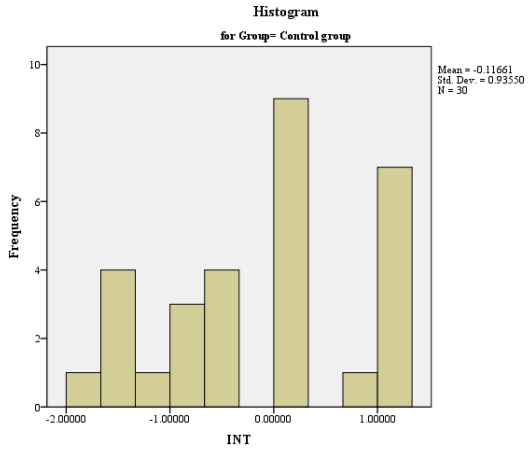

Histogram
for Group= Game group
Mean = 0.02397
Std. Dev. = 1.08403
N = 22
Frequency
EM

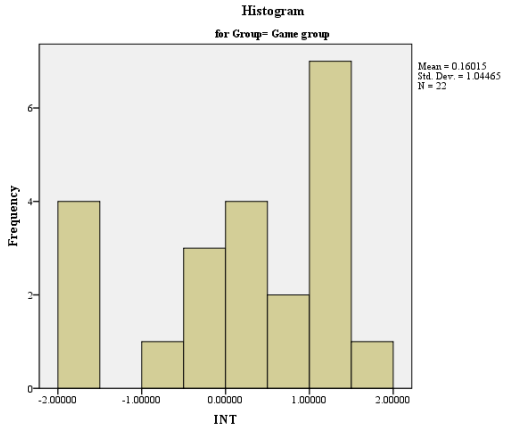

Normal Q-Q Plot of EM
for Group= Control group
Expected Normal
Observed Value

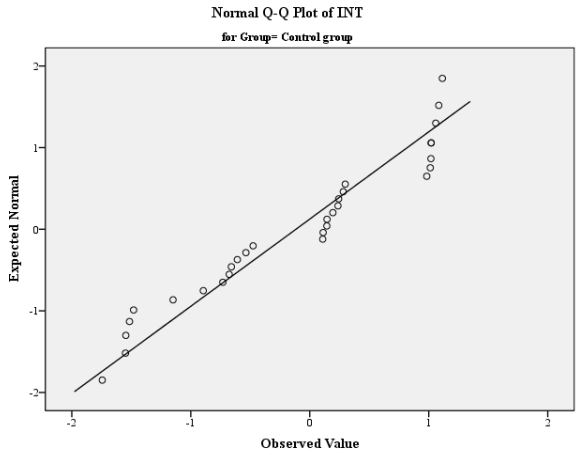

Normal Q-Q Plot of EM
for Group= Game group
Expected Normal
Observed Value

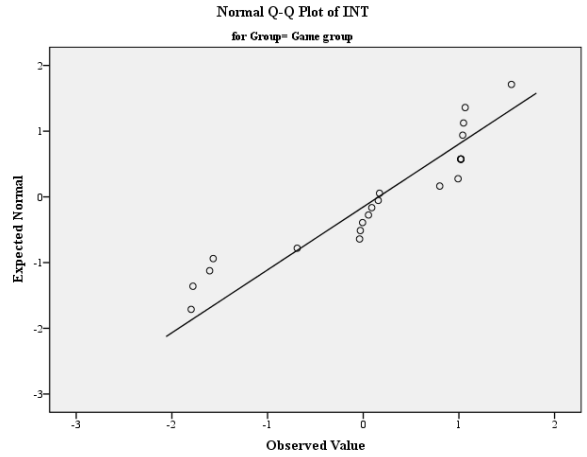

EM
64
342
180
277
207
Control group
Game group
Group

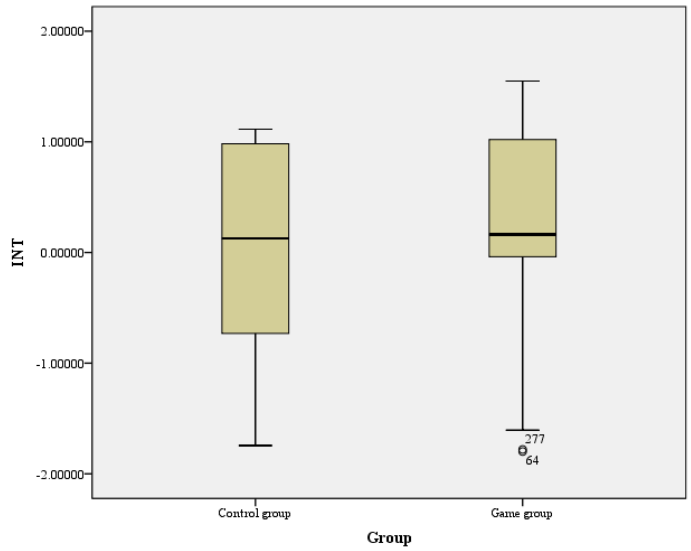
SG vs INT


Histogram
for Group= Control group
Mean = -0.11661
Std. Dev. = 0.93550
N = 30
Frequency
INT

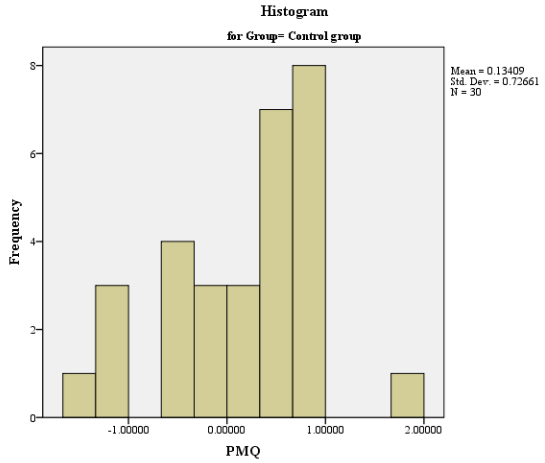

Histogram
for Group= Game group
Mean = 0.16015
Std. Dev. = 1.04465
N = 22
Frequency
INT

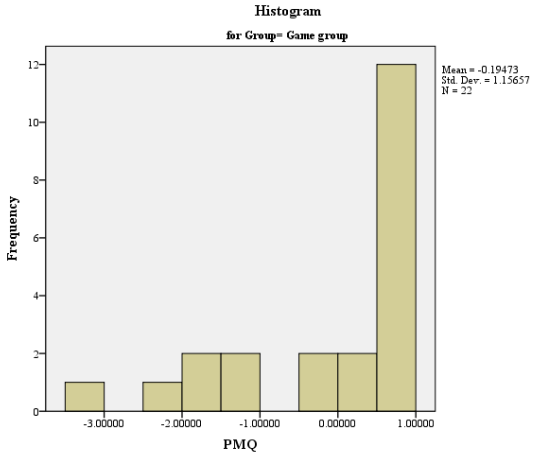

Normal Q-Q Plot of INT
for Group= Control group
Expected Normal
Observed Value

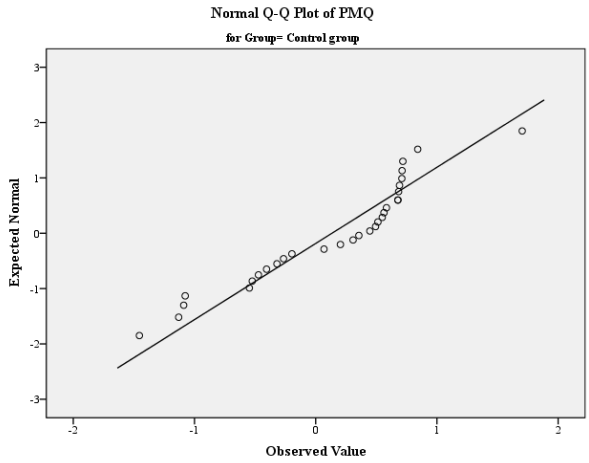

Normal Q-Q Plot of INT
for Group= Game group
Expected Normal
Observed Value

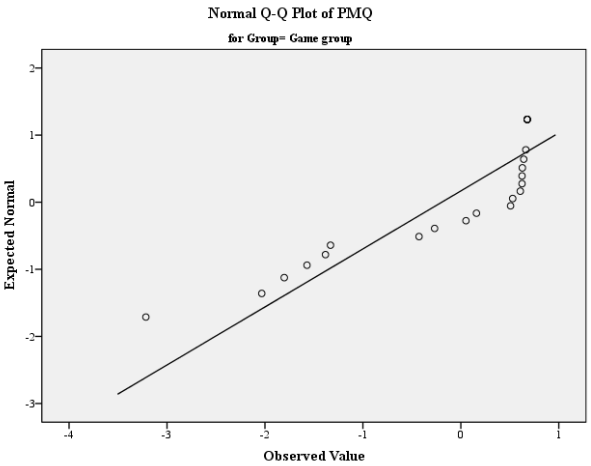

INT
Control group
Game group
Group
277
64

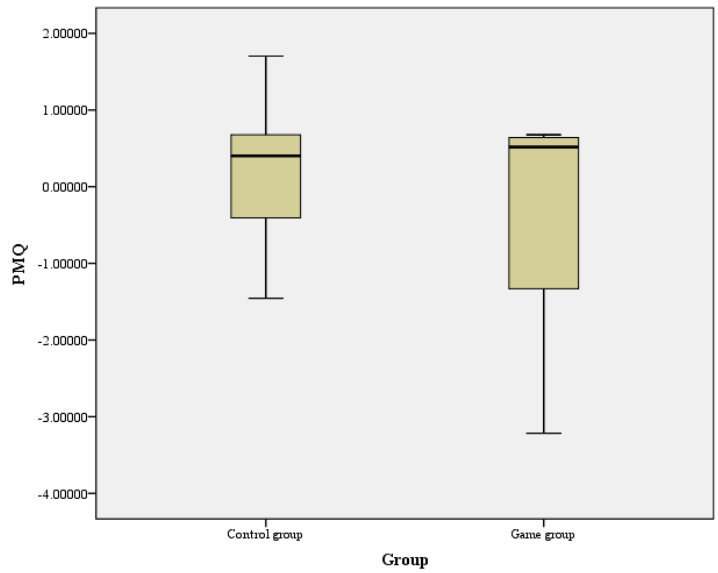
SG vs PMQ


Histogram
for Group= Control group
Mean = 0.13409
Std. Dev. = 0.72661
N = 30
Frequency
PMQ


Histogram
for Group= Game group
Mean = -0.19473
Std. Dev. = 1.15657
N = 22
Frequency
PMQ

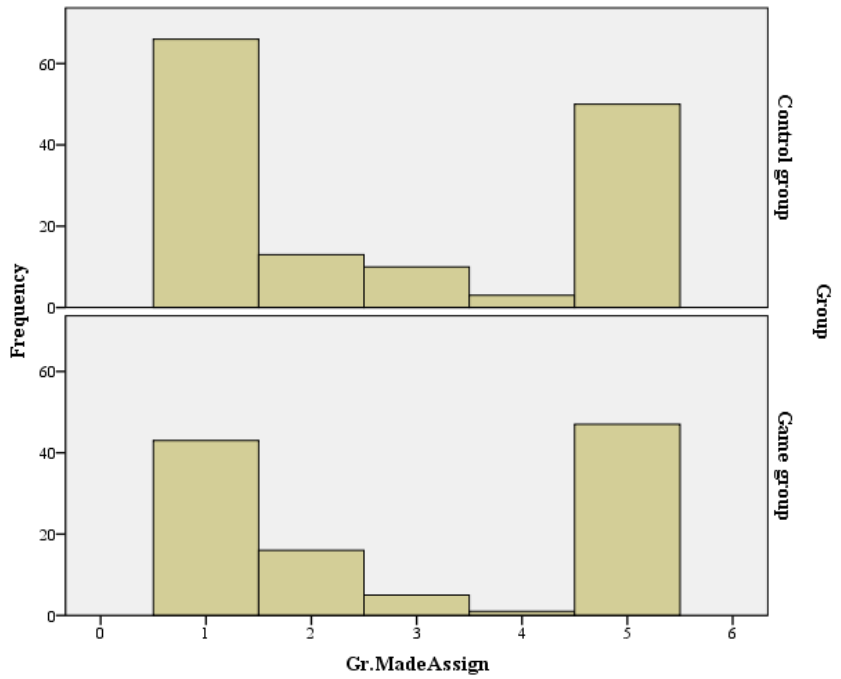

Normal Q-Q Plot of PMQ
for Group= Control group
Expected Normal
Observed Value

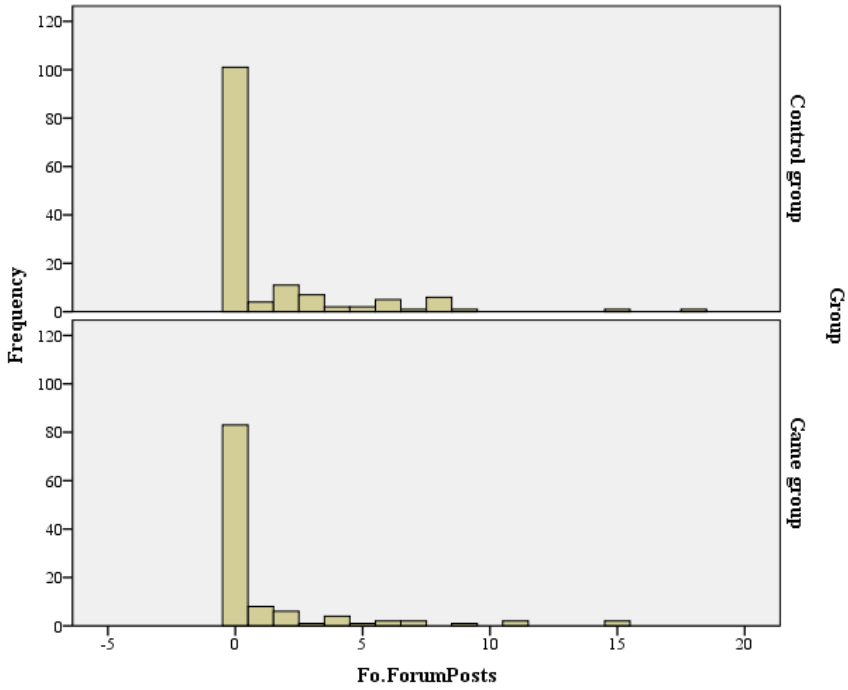

Normal Q-Q Plot of PMQ
for Group= Game group
Expected Normal
Observed Value

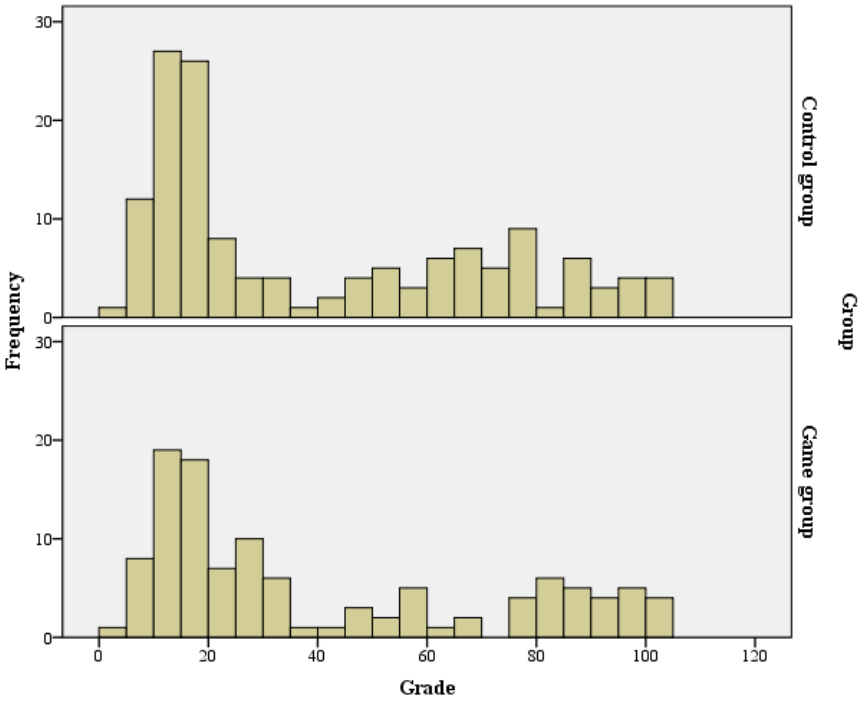

PMQ
Control group
Game group
Group

Appendix 13: Part 3: Visual normality check for number of assignments made, forum posts, and grade

This appendix shows the histograms used see if the number of assignments made, forums posts and grade were approximately normally distributed for the control group and game group.






Frequency
30
20
10
0
Control group
30
20
10
0
Game group
Group
0
20
40
60
80
100
120
Grade

Appendix 14: Supportive results DelftX questions resembling factors

This appendix shows the test results when analysing the answers to the questions in the DelftX mid-questionnaire and post-questionnaire that somewhat resembled the questions used to measure the factors in the additional questionnaires. By analysing these results, it can be seen if any other results are obtained and these results can be used to provide context for the main statistical analysis. This also indicates to some extent the reliability of the additional questionnaire, as it gives an idea about the consistency of the respondents. This appendix contains all analyses regarding the DelftX questions, for bot theoretical framework part 1 and part 2.

Intrinsic motivation

Question: Rank your current course enjoyment (1, 2, 3, 4, 5).

Source: DelftX mid-questionnaire

Results: No significant differences.

Intrinsic motivation vs completion

Group Statistics

	Gr.EligCert	N	Mean	Std. Deviation	Std. Error Mean
Course enjoyment	Not eligable (not completed)	10	4.00	0.943	0.298
	Eligable (completed)	20	4.15	0.813	0.182

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Course enjoyment	Equal variances assumed	0.548	0.465	-0.452	28	0.655	-0.150	0.332	-0.830	0.530
	Equal variances not assumed			-0.430	15.891	0.673	-0.150	0.349	-0.891	0.591

Intrinsic motivation vs grade & number of assignments made

Correlations

		tmq.enjoy	Grade	Number of assignments made
tmq.enjoy	Pearson Correlation	1	0.179	0.233
	Sig. (2-tailed)		0.344	0.215
	N	30	30	30
Grade	Pearson Correlation	0.179	1	0.917**
	Sig. (2-tailed)	0.344		0.000
	N	30	134	134
Number of assignments made	Pearson Correlation	0.233	0.917**	1
	Sig. (2-tailed)	0.215	0.000	
	N	30	134	134

**. Correlation is significant at the 0.01 level (2-tailed).

Intrinsic motivation vs groups

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tmq.enjoy	Control group	21	4.10	0.889	0.194
	Game group	9	4.11	0.782	0.261

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tmq.enjoy	Equal variances assumed	1.203	0.282	-0.046	28	0.963	-0.016	0.343	-0.718	0.686
	Equal variances not assumed			-0.049	17.212	0.962	-0.016	0.325	-0.701	0.669

Extrinsic motivation

Source: DelftX post-questionnaire

Results: No significant differences

Questions:

- Tpoq.em.app = I can apply what I have learned in this course in my daily work
- Tpoq.em.careeradv = This course help me advance in my career / professional field

Extrinsic motivation vs completion

Group Statistics

	Gr.EligCert	N	Mean	Std. Deviation	Std. Error Mean
tpoq.em.appl	Not eligable (not completed)	5	3.40	1.342	0.600
	Eligable (completed)	15	4.00	0.655	0.169
tpoq.em.careeradv	Not eligable (not completed)	5	3.40	1.342	0.600
	Eligable (completed)	15	3.67	0.724	0.187

Independent Samples Test

		Levene's Test for Equality of Variances		**t-test for Equality of Means**						
									95% Confidence Interval of the Difference	
		F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**Lower**	**Upper**
tpoq.em.appl	Equal variances assumed	3.419	0.081	-1.357	18	0.192	-0.600	0.442	-1.529	0.329
	Equal variances not assumed			-0.963	4.652	0.383	-0.600	0.623	-2.239	1.039
tpoq.em.careeradv	Equal variances assumed	1.870	0.188	-0.575	18	0.573	-0.267	0.464	-1.242	0.708
	Equal variances not assumed			-0.424	4.801	0.690	-0.267	0.628	-1.902	1.369

Extrinsic motivation vs grade & assignments made

Correlations

		Grade	Number of assignments made	tpoq.em.appl	tpoq.em.careeradv
Grade	Pearson Correlation	1	0.917**	0.194	0.080
	Sig. (2-tailed)		0.000	0.412	0.738
	N	134	134	20	20
Number of assignments made	Pearson Correlation	0.917**	1	-0.040	-0.107
	Sig. (2-tailed)	0.000		0.866	0.654
	N	134	134	20	20
tpoq.em.appl	Pearson Correlation	0.194	-0.040	1	0.736**
	Sig. (2-tailed)	0.412	0.866		0.000
	N	20	20	20	20
tpoq.em.careeradv	Pearson Correlation	0.080	-0.107	0.736**	1
	Sig. (2-tailed)	0.738	0.654	0.000	
	N	20	20	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Extrinsic motivation vs groups

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tpoq.em.appl	Control group	13	4.08	0.641	0.178
	Game group	7	3.43	1.134	0.429
tpoq.em.careeradv	Control group	13	3.69	0.751	0.208
	Game group	7	3.43	1.134	0.429

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.em.appl	Equal variances assumed	2.223	0.153	1.651	18	0.116	0.648	0.393	-0.177	1.474
	Equal variances not assumed			1.398	8.119	0.199	0.648	0.464	-0.419	1.715
tpoq.em.careeradv	Equal variances assumed	0.571	0.460	0.627	18	0.538	0.264	0.421	-0.620	1.147
	Equal variances not assumed			0.553	8.921	0.594	0.264	0.477	-0.816	1.343

Interactivity

Source: DelftX post-questionnaire

Results: No significant differences

Questions:

- Tpoq.int.act.forum = The instructor and / or teaching assistants had an active role on the course forums.
- Tpoq.int.respond.forum = The instructor and / or teaching assistants adequately responded to questions from students.
- Tpoq.int.mot = The teachers motivated me to follow the course
- Tpoq.int.discuss = On the course's discussion forum I felt encouraged to discuss the course with other students
- Tpoq.int.tohelp = I felt encouraged to help other students with their course work
- Tpoq.int.askstudents = I felt encourage to ask for help from other students
- Tpoq.int.askteachers = I felt encouraged to ask for help from the course instructors
- Tpoq.int.questionansw = the questions that I asked were answered
- Tpoq.int.answer.useful = the answers that my questions received were useful

Interactivity vs completion

Group Statistics

	Gr.EligCert	N	Mean	Std. Deviation	Std. Error Mean
tpoq.int.act.forum	Not eligable (not completed)	5	3.00	1.225	0.548
	Eligable (completed)	15	3.27	0.799	0.206
tpoq.int.respond.forum	Not eligable (not completed)	5	3.20	1.304	0.583
	Eligable (completed)	15	3.13	0.834	0.215
tpoq.int.mot	Not eligable (not completed)	6	2.83	1.472	0.601
	Eligable (completed)	15	3.27	0.961	0.248
tpoq.int.dicuss	Not eligable (not completed)	6	3.17	1.169	0.477
	Eligable (completed)	15	3.20	0.414	0.107
tpoq.int.tohelp	Not eligable (not completed)	6	3.33	1.211	0.494
	Eligable (completed)	15	3.27	0.458	0.118
tpoq.int.askstudents	Not eligable (not completed)	6	3.00	1.095	0.447
	Eligable (completed)	14	3.21	0.426	0.114
tpoq.int.askteacher	Not eligable (not completed)	6	3.17	1.169	0.477
	Eligable (completed)	15	3.40	0.632	0.163
tpoq.int.questanswer	Not eligable (not completed)	6	3.00	1.095	0.447
	Eligable (completed)	15	3.33	0.617	0.159
tpoq.int.answer.usef	Not eligable (not completed)	6	3.50	1.378	0.563
	Eligable (completed)	15	3.47	0.516	0.133

Independent Samples Test

									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.int.act.forum	Equal variances assumed	0.468	0.503	-0.567	18	0.578	-0.267	0.470	-1.255	0.722
	Equal variances not assumed			-0.456	5.185	0.667	-0.267	0.585	-1.755	1.222
tpoq.int.respond.forum	Equal variances assumed	1.425	0.248	0.135	18	0.894	0.067	0.495	-0.973	1.106
	Equal variances not assumed			0.107	5.138	0.919	0.067	0.622	-1.518	1.652
tpoq.int.mot	Equal variances assumed	2.802	0.111	-0.802	19	0.432	-0.433	0.540	-1.564	0.697
	Equal variances not assumed			-0.667	6.780	0.527	-0.433	0.650	-1.981	1.114
tpoq.int.dicuss	Equal variances assumed	6.077	0.023	-0.099	19	0.922	-0.033	0.337	-0.738	0.671
	Equal variances not assumed			-0.068	5.509	0.948	-0.033	0.489	-1.256	1.190
tpoq.int.tohelp	Equal variances assumed	6.244	0.022	0.188	19	0.853	0.067	0.355	-0.677	0.810
	Equal variances not assumed			0.131	5.581	0.900	0.067	0.508	-1.200	1.334
tpoq.int.askstudents	Equal variances assumed	2.006	0.174	-0.645	18	0.527	-0.214	0.332	-0.913	0.484

	Equal variances not assumed			-0.464	5.659	0.660	-0.214	0.461	-1.360	0.932
tpoq.int.askteacher	Equal variances assumed	1.708	0.207	-0.597	19	0.557	-0.233	0.391	-1.051	0.585
	Equal variances not assumed			-0.463	6.209	0.659	-0.233	0.504	-1.458	0.991
tpoq.int.questanswer	Equal variances assumed	0.329	0.573	-0.893	19	0.383	-0.333	0.373	-1.114	0.448
	Equal variances not assumed			-0.702	6.314	0.508	-0.333	0.475	-1.481	0.814
tpoq.int.answer.usef	Equal variances assumed	5.840	0.026	0.083	19	0.935	0.033	0.403	-0.810	0.877
	Equal variances not assumed			0.058	5.571	0.956	0.033	0.578	-1.409	1.475

Interactivity vs grade & assignments made

		Grade	Number of assignments made
tpoq.int.act.forum	Pearson Correlation	0.221	0.072
	Sig. (2-tailed)	0.349	0.762
	N	20	20
tpoq.int.respond.forum	Pearson Correlation	0.062	-0.179
	Sig. (2-tailed)	0.795	0.451
	N	20	20
tpoq.int.mot	Pearson Correlation	0.411	0.418
	Sig. (2-tailed)	0.064	0.059
	N	21	21
tpoq.int.dicuss	Pearson Correlation	0.183	0.088
	Sig. (2-tailed)	0.428	0.705
	N	21	21
tpoq.int.tohelp	Pearson Correlation	0.007	-0.167
	Sig. (2-tailed)	0.974	0.470
	N	21	21
tpoq.int.askstudents	Pearson Correlation	0.220	0.072
	Sig. (2-tailed)	0.352	0.762
	N	20	20
tpoq.int.askteacher	Pearson Correlation	0.212	-0.131
	Sig. (2-tailed)	0.356	0.570
	N	21	21
tpoq.int.questanswer	Pearson Correlation	0.362	0.097
	Sig. (2-tailed)	0.107	0.675
	N	21	21
tpoq.int.answer.usef	Pearson Correlation	0.031	-0.331
	Sig. (2-tailed)	0.894	0.143
	N	21	21
Grade	Pearson Correlation	1	0.917**
	Sig. (2-tailed)		0.000
	N	134	134
Number of assignments made	Pearson Correlation	0.917**	1
	Sig. (2-tailed)	0.000	
	N	134	134

Interactivity vs groups

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tpoq.int.act.forum	Control group	13	3.23	0.832	0.231
	Game group	7	3.14	1.069	0.404
tpoq.int.respond.for	Control group	13	3.08	0.862	0.239
um	Game group	7	3.29	1.113	0.421
tpoq.int.mot	Control group	13	3.31	0.947	0.263
	Game group	8	2.88	1.356	0.479
tpoq.int.dicuss	Control group	13	3.23	0.439	0.122
	Game group	8	3.13	0.991	0.350
tpoq.int.tohelp	Control group	13	3.31	0.480	0.133
	Game group	8	3.25	1.035	0.366
tpoq.int.askstudents	Control group	12	3.25	0.452	0.131
	Game group	8	3.00	0.926	0.327
tpoq.int.askteacher	Control group	13	3.38	0.650	0.180
	Game group	8	3.25	1.035	0.366
tpoq.int.questanswe	Control group	13	3.31	0.630	0.175
r	Game group	8	3.13	0.991	0.350
tpoq.int.answer.usef	Control group	13	3.46	0.519	0.144
	Game group	8	3.50	1.195	0.423

Independent Samples Test

									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.int.act.forum	Equal variances assumed	0.245	0.627	0.204	18	0.840	0.088	0.430	-0.816	0.992
	Equal variances not assumed			0.189	10.020	0.854	0.088	0.465	-0.949	1.124
tpoq.int.respond.forum	Equal variances assumed	0.674	0.422	-0.467	18	0.646	-0.209	0.447	-1.148	0.730
	Equal variances not assumed			-0.432	9.986	0.675	-0.209	0.484	-1.287	0.869
tpoq.int.mot	Equal variances assumed	2.773	0.112	0.863	19	0.399	0.433	0.501	-0.616	1.482
	Equal variances not assumed			0.791	11.243	0.445	0.433	0.547	-0.768	1.633
tpoq.int.dicuss	Equal variances assumed	2.083	0.165	0.339	19	0.739	0.106	0.312	-0.548	0.760
	Equal variances not assumed			0.285	8.715	0.782	0.106	0.371	-0.737	0.949
tpoq.int.tohelp	Equal variances assumed	2.896	0.105	0.175	19	0.863	0.058	0.330	-0.634	0.749
	Equal variances not assumed			0.148	8.888	0.886	0.058	0.389	-0.825	0.940
tpoq.int.askstudents	Equal variances assumed	0.296	0.593	0.809	18	0.429	0.250	0.309	-0.399	0.899

	Equal variances not assumed			0.709	9.255	0.496	0.250	0.352	-0.544	1.044
tpoq.int.asktheacher	Equal variances assumed	0.802	0.382	0.368	19	0.717	0.135	0.366	-0.631	0.900
	Equal variances not assumed			0.330	10.455	0.748	0.135	0.408	-0.769	1.038
tpoq.int.questanswer	Equal variances assumed	0.319	0.579	0.519	19	0.610	0.183	0.352	-0.554	0.919
	Equal variances not assumed			0.467	10.539	0.650	0.183	0.392	-0.684	1.049
tpoq.int.answer.usef	Equal variances assumed	3.452	0.079	-0.103	19	0.919	-0.038	0.375	-0.823	0.746
	Equal variances not assumed			-0.086	8.650	0.933	-0.038	0.446	-1.055	0.978

Perceived MOOC Quality

Source: DelftX post-questionnaire

Results: No significant differences

Questions:

- Tpoq.pmq.learnlot = I learned a lot from this course
- Tpoq.pmq.worthmytime = This course was well worth my time
- Tpoq.pmq.recomm = I would recommend this course to another person
- Tpoq.pmq.diff = How would you rate the difficulty level of this course?
- Tpoq.pmq.workreq = How would you describe the amount of work required for the course?
- Tpoq.pmq.durat = How would you describe the duration (in weeks) of the course?

PMQ vs completion

Group Statistics

	Gr.EligCert	N	Mean	Std. Deviation	Std. Error Mean
tpoq.pmq.learnlot	Not eligable (not completed)	5	3.40	1.342	0.600
	Eligable (completed)	15	4.13	0.640	0.165
tpoq.pmq.worthmytime	Not eligable (not completed)	5	3.40	1.342	0.600
	Eligable (completed)	15	4.00	0.378	0.098
tpoq.pmq.reccom	Not eligable (not completed)	5	3.40	1.342	0.600
	Eligable (completed)	15	4.20	0.414	0.107
tpoq.pmq.course.diff	Not eligable (not completed)	5	2.80	0.447	0.200
	Eligable (completed)	15	2.60	0.507	0.131
tpoq.pmq.course.workreq	Not eligable (not completed)	6	3.17	0.408	0.167
	Eligable (completed)	14	3.07	0.730	0.195
tpoq.pmq.course.durat	Not eligable (not completed)	6	3.33	0.516	0.211
	Eligable (completed)	14	3.07	0.616	0.165

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
								95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.pmq.learnlot	Equal variances assumed	3.266	0.087	-1.675	18	0.111	-0.733	0.438	-1.653	0.186
	Equal variances not assumed			-1.178	4.622	0.296	-0.733	0.622	-2.373	0.907
tpoq.pmq.worthmytime	Equal variances assumed	10.665	0.004	-1.625	18	0.121	-0.600	0.369	-1.376	0.176
	Equal variances not assumed			-0.987	4.214	0.377	-0.600	0.608	-2.255	1.055
tpoq.pmq.recom	Equal variances assumed	8.000	0.011	-2.121	18	0.048	-0.800	0.377	-1.592	-0.008
	Equal variances not assumed			-1.313	4.257	0.256	-0.800	0.609	-2.453	0.853
tpoq.pmq.course.diff	Equal variances assumed	4.000	0.061	0.783	18	0.444	0.200	0.255	-0.336	0.736
	Equal variances not assumed			0.837	7.756	0.428	0.200	0.239	-0.354	0.754
tpoq.pmq.course.workreq	Equal variances assumed	1.439	0.246	0.297	18	0.770	0.095	0.320	-0.578	0.768
	Equal variances not assumed			0.371	16.312	0.715	0.095	0.257	-0.448	0.638

tpoq.pmq.course.durat	Equal variances assumed	0.057	0.814	0.910	18	0.375	0.262	0.288	-0.343	0.867
	Equal variances not assumed			0.979	11.331	0.348	0.262	0.267	-0.325	0.848

PMQ vs grade & assignments made

Correlations

		Grade	Number of assignments made	tpoq.pmq.learnlot	tpoq.pmq.worthmyti me	tpoq.pmq.reccom	tpoq.pmq.course.diff	tpoq.pmq.course.workreq	tpoq.pmq.course.durat
Grade	Pearson Correlation	1	0.917**	0.234	0.293	0.341	0.043	-0.043	-0.198
	Sig. (2-tailed)		0.000	0.322	0.209	0.141	0.858	0.856	0.402
	N	134	134	20	20	20	20	20	20
Number of assignments made	Pearson Correlation	0.917*	1	-0.013	-0.047	0.000	0.313	-0.285	-0.284
	Sig. (2-tailed)	0.000		0.956	0.843	1.000	0.180	0.223	0.225
	N	134	134	20	20	20	20	20	20

PMQ vs groups

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tpoq.pmq.learnlot	Control group	13	4.15	0.555	0.154
	Game group	7	3.57	1.272	0.481
tpoq.pmq.worthmyti me	Control group	13	4.08	0.277	0.077
	Game group	7	3.43	1.134	0.429
tpoq.pmq.reccom	Control group	13	4.23	0.439	0.122
	Game group	7	3.57	1.134	0.429
tpoq.pmq.course.dif f	Control group	12	2.58	0.515	0.149
	Game group	8	2.75	0.463	0.164
tpoq.pmq.course.wo rkreq	Control group	12	3.00	0.739	0.213
	Game group	8	3.25	0.463	0.164
tpoq.pmq.course.du rat	Control group	12	3.00	0.603	0.174
	Game group	8	3.38	0.518	0.183

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
								95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.pmq.learnlot	Equal variances assumed	3.649	0.072	1.440	18	0.167	0.582	0.405	-0.268	1.432
	Equal variances not assumed			1.153	7.253	0.285	0.582	0.505	-0.603	1.768
tpoq.pmq.worthmytime	Equal variances assumed	10.036	0.005	1.996	18	0.061	0.648	0.325	-0.034	1.331
	Equal variances not assumed			1.489	6.389	0.184	0.648	0.435	-0.402	1.698
tpoq.pmq.recom	Equal variances assumed	2.564	0.127	1.885	18	0.076	0.659	0.350	-0.076	1.394
	Equal variances not assumed			1.480	6.983	0.183	0.659	0.445	-0.395	1.713
tpoq.pmq.course.diff	Equal variances assumed	2.339	0.144	-0.737	18	0.471	-0.167	0.226	-0.642	0.308
	Equal variances not assumed			-0.754	8	0.462	-0.167	0.221	-0.635	0.301
tpoq.pmq.course.workreq	Equal variances assumed	0.400	0.535	-0.849	18	0.407	-0.250	0.295	-0.869	0.369
	Equal variances not assumed			-0.930	5	0.365	-0.250	0.269	-0.815	0.315

tpoq.pmq.course.durat	Equal variances assumed	0.569	0.460	-1.438	18	0.168	-0.375	0.261	-0.923	0.173
	Equal variances not assumed			-1.4850	16.70	0.156	-0.375	0.253	-0.909	0.159

Appendix 15: Demographics vs groups

This appendix shows the results of the tests conducted to assess if the groups were equal in terms of student demographics. The results show that the groups only differed in the average age of the students.

Groups vs Paying students

Group * EnrollTrack Crosstabulation

			EnrollTrack		
			Not payed	**Payed**	**Total**
Group	Control group	Count	72	7	79
		% within Group	91.1%	8.9%	100.0%
	Game group	Count	52	3	55
		% within Group	94.5%	5.5%	100.0%
Total		Count	124	10	134
		% within Group	92.5%	7.5%	100.0%

Chi-Square Tests

	Value	**df**	**Asymptotic Significance (2-sided)**	**Exact Sig. (2-sided)**	**Exact Sig. (1-sided)**
Pearson Chi-Square	0.545[a]	1	0.460		
Continuity Correction[b]	0.163	1	0.686		
Likelihood Ratio	0.564	1	0.453		
Fisher's Exact Test				0.525	0.350
Linear-by-Linear Association	0.541	1	0.462		
N of Valid Cases	134				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.10.

b. Computed only for a 2x2 table

Groups vs. Age & English proficiency

Age was found to be significantly different between groups. Students in game group were approximately 7 years older.

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tpeq.age	Control group	53	33.26	12.935	1.777
	Game group	36	40.11	13.813	2.302
tpeq.engprof	Control group	55	4.02	1.045	0.141
	Game group	37	4.35	0.716	0.118

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
								95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpeq.age	Equal variances assumed	2.364	0.128	-2.385	87	0.019	-6.847	2.871	-12.554	-1.140
	Equal variances not assumed			-2.355	71.936	0.021	-6.847	2.908	-12.644	-1.050
tpeq.engprof	Equal variances assumed	1.410	0.238	-1.690	90	0.095	-0.333	0.197	-0.725	0.059
	Equal variances not assumed			-1.815	89.957	0.073	-0.333	0.184	-0.698	0.032

Groups vs gender

Group * tpeq.gender Crosstabulation

			tpeq.gender		
			Male	**Female**	**Total**
Group	Control group	Count	33	21	54
		% within Group	61.1%	38.9%	100.0%
	Game group	Count	24	12	36
		% within Group	66.7%	33.3%	100.0%
Total		Count	57	33	90
		% within Group	63.3%	36.7%	100.0%

Chi-Square Tests

	Value	**df**	**Asymptotic Significance (2-sided)**	**Exact Sig. (2-sided)**	**Exact Sig. (1-sided)**
Pearson Chi-Square	0.287[a]	1	0.592		
Continuity Correction[b]	0.098	1	0.755		
Likelihood Ratio	0.289	1	0.591		
Fisher's Exact Test				0.659	0.379
Linear-by-Linear Association	0.284	1	0.594		
N of Valid Cases	90				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.20.

b. Computed only for a 2x2 table

Groups vs degree

Crosstab

			tpeq.degree			
			Professional degree or lower	**Bachelor's degree**	**Master's degree or higher**	Total
Group	Control group	Count	12	19	20	51
		% within Group	23.5%	37.3%	39.2%	100.0%
	Game group	Count	6	14	15	35
		% within Group	17.1%	40.0%	42.9%	100.0%
Total		Count	18	33	35	86
		% within Group	20.9%	38.4%	40.7%	100.0%

Chi-Square Tests

	Value	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	0.513[a]	2	0.774
Likelihood Ratio	0.522	2	0.770
Linear-by-Linear Association	0.357	1	0.550
N of Valid Cases	86		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.33.

Groups vs current occupation

Crosstab

			tpeq.occupt							
			Student	Recently graduated	Working professional	Looking for a job	Parent / caregiver	Retired	Other	Total
Group	Control group	Count	8	1	37	5	1	1	0	53
		% within Group	15.1%	1.9%	69.8%	9.4%	1.9%	1.9%	0.0%	100.0%
	Game group	Count	4	0	25	2	0	3	2	36
		% within Group	11.1%	0.0%	69.4%	5.6%	0.0%	8.3%	5.6%	100.0%
Total		Count	12	1	62	7	1	4	2	89
		% within Group	13.5%	1.1%	69.7%	7.9%	1.1%	4.5%	2.2%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.948[a]	6	0.326
Likelihood Ratio	8.349	6	0.214
Linear-by-Linear Association	2.827	1	0.093
N of Valid Cases	89		

a. 11 cells (78.6%) have expected count less than 5. The minimum expected count is 0.40.

Groups vs continent of origin

Crosstab

			tpeq.continent						
			Europe	North America	South America	Africa	Asia	Australia&Oceania	Total
Group	Control group	Count	21	8	5	6	12	2	54
		% within Group	38.9%	14.8%	9.3%	11.1%	22.2%	3.7%	100.0%
	Game group	Count	11	6	4	2	13	0	36
		% within Group	30.6%	16.7%	11.1%	5.6%	36.1%	0.0%	100.0%
Total		Count	32	14	9	8	25	2	90
		% within Group	35.6%	15.6%	10.0%	8.9%	27.8%	2.2%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.127[a]	5	0.531
Likelihood Ratio	4.857	5	0.434
Linear-by-Linear Association	0.486	1	0.486
N of Valid Cases	90		

a. 5 cells (41.7%) have expected count less than 5. The minimum expected count is 0.80.

Groups vs enrolment intention

Group * tpeq.intention

Group * tpeq.intention Crosstabulation

			tpeq.intention		
			I want to complete this course	I do not intend to complete this course	Total
Group	Control group	Count	47	8	55
		% within Group	85.5%	14.5%	100.0%
	Game group	Count	33	4	37
		% within Group	89.2%	10.8%	100.0%
Total		Count	80	12	92
		% within Group	87.0%	13.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	0.272[a]	1	0.602		
Continuity Correction[b]	0.042	1	0.837		
Likelihood Ratio	0.277	1	0.598		
Fisher's Exact Test				0.756	0.425
Linear-by-Linear Association	0.269	1	0.604		
N of Valid Cases	92				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.83.

b. Computed only for a 2x2 table

Groups vs enrolment motivation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Group tpeq.enrolmot	*91	67.9%	43	32.1%	134	100.0%

Group * tpeq.enrolmot Crosstabulation

			tpeq.enrolmot		
			To use in daily work activities or improve career prospects	Interested in subject and want to learn more	Total
Group	Control group	Count	36	18	54
		% within tpeq.enrolmot	65.5%	50.0%	59.3%
	Game group	Count	19	18	37
		% within tpeq.enrolmot	34.5%	50.0%	40.7%
Total		Count	55	36	91
		% within tpeq.enrolmot	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.154[a]	1	0.142		
Continuity Correction[b]	1.561	1	0.212		
Likelihood Ratio	2.147	1	0.143		
Fisher's Exact Test				0.191	0.106
Linear-by-Linear Association	2.130	1	0.144		
N of Valid Cases	91				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 14.64.

b. Computed only for a 2x2 table

Appendix 16: Quantitative serious game evaluation DelftX post-questionnaire

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
tpoq.pmq.assclear	Control group	12	3.08	0.900	0.260
	Game group	8	3.88	1.246	0.441
tpoq.pmq.asseasy	Control group	11	2.45	0.934	0.282
	Game group	8	3.00	1.195	0.423
tpoq.pmq.assdiff	Control group	12	3.42	0.793	0.229
	Game group	8	3.00	1.195	0.423
tpoq.pmq.assreflcontent	Control group	12	3.33	1.073	0.310
	Game group	8	3.50	1.195	0.423
tpoq.pmq.assmadeapply	Control group	12	3.92	0.515	0.149
	Game group	8	3.63	1.188	0.420
tpoq.pmq.assmade.iden	Control group	12	3.92	0.515	0.149
	Game group	8	3.50	1.195	0.423

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
								95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpoq.pmq.assclear	Equal variances assumed	0.020	0.888	-1.654	18	0.115	-0.792	0.479	-1.797	0.214
	Equal variances not assumed			-1.5478	11.80	0.148	-0.792	0.512	-1.908	0.325
tpoq.pmq.asseasy	Equal variances assumed	1.035	0.323	-1.118	17	0.279	-0.545	0.488	-1.574	0.483

	Equal variances not assumed			-1.0749	12.82	0.303	-0.545	0.508	-1.644	0.553
tpoq.pmq.assdiff	Equal variances assumed	2.739	0.115	0.942	18	0.359	0.417	0.442	-0.513	1.346
	Equal variances not assumed			0.867	11.102	0.404	0.417	0.481	-0.640	1.473
tpoq.pmq.assreflcontent	Equal variances assumed	0.002	0.962	-0.325	18	0.749	-0.167	0.512	-1.243	0.909
	Equal variances not assumed			-0.3186	13.97	0.755	-0.167	0.524	-1.291	0.957
tpoq.pmq.assmadeapply	Equal variances assumed	3.475	0.079	0.758	18	0.458	0.292	0.385	-0.517	1.100
	Equal variances not assumed			0.655	8.776	0.529	0.292	0.445	-0.720	1.303
tpoq.pmq.assmade.iden	Equal variances assumed	4.940	0.039	1.078	18	0.295	0.417	0.387	-0.396	1.229
	Equal variances not assumed			0.930	8.754	0.377	0.417	0.448	-0.601	1.434

Appendix 17: Qualitative game evaluation results

User-id	Source of qualitative answer
	Forum (only responses for game 1)
	Please discuss your experiences with the game here.
-	It was a good experience, but I think the reference of some problems may differ from one country to another. But it was nice
-	It was a good experience, where I was able to test my speed and problem solving skills.
-	I like the game agree that the problem owner may differ in different context.
-	Really good practice, but I don't like speed part of the game.
-	It was interesting, but agreed with (…) I didn't like the speed part of the game. I felt too rushed and it was harder to concentrate.
-	Time definitely puts pressure but the game is great forces you to think at what level are the actors
-	Couldn't answer of any problem statement. Don't know what was the problem :-(
	Mid questionnaire
	Which aspect of this course do you especially like? Please name one. (13 responses from game group)
6510	Games are fun, the reading material is great
5828	I liked that I am able to test my knowledge though the use of quiz, test and diagramming.
11499	Videos and quiz
	Which aspect of this course would you like us to change or improve? Please name one.
13121	Time for choosing the answers in games may be increased slightly.
8852	The timing on the quiz (the timing is too short)
11499	Less games
2626	The way quiz are designed.
	Additional questionnaire
	If you have any improvements, praise, or other comments regarding the games please feel free to share them below
14728	Good
10089	I never knew about them
969	Improve on the timing of games please.
17194	The weak point of the games is that non-native English speakers usually require more time to process the question and options for an answer. It would be beneficial to adapt the time given to the learners to their level of English proficiency.
	Post questionnaire
	Which one aspect of this course did you like the most? Please explain. (3 Reponses from game group)
6510	Games. Like I said, they are great for practice and at the same time they are fun.

Appendix 18: Demographics vs three student retention variables

A supportive analysis was conducted to analyse the influence of the student demographics on the three student retention variables. It was found that the completion rate was higher for students who paid to receive a certificate at some point during the course. Furthermore, paying students were 38% ($p < 0.002$) more likely to finish the course, score a 21% higher grade and made 1.1 more assignments. Completers were found to be 6 years younger than non-completers, 32 and 28 respectively ($p < 0.046$). Age was also significantly negatively correlated with the number of assignments made ($p<0.045$; $r = -0.213$) and the grade ($p < 0.033$; $r = -0.23$). Students who from the beginning (Delft pre-questionnaire) indicated they intended to complete the course, made 1.3 more assignments ($p < 0.012$) and obtained a 23% higher grade ($p < 0.013$). Lastly, a significant ($p<0.014$) difference (-14%) between males and females was found for the average grade obtained, but no significant differences in completion or number of assignments made. Lastly, no significant differences were found for completion, assignments made and grade regarding the other demographic variables: English proficiency, highest obtained degree, current occupation, continent of origin, and enrolment motivation.

Contents

Completion vs demographics

Paying students

Significant difference ($p < 0.002$) of 38% more students who payed finished the course.

EnrollTrack * Gr.EligCert Crosstabulation

			Gr.EligCert Eligible (completed)	Not eligible	Total
EnrollTrack	Not paying students	Count	62	175	237
		% within EnrollTrack	26.2%	73.8%	100.0%
	Paying students	Count	11	6	17
		% within EnrollTrack	64.7%	35.3%	100.0%
Total		Count	73	181	254
		% within EnrollTrack	28.7%	71.3%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.507[a]	1	0.001		
Continuity Correction[b]	9.702	1	0.002		
Likelihood Ratio	10.208	1	0.001		
Fisher's Exact Test				0.002	0.002
Linear-by-Linear Association	11.462	1	0.001		
N of Valid Cases	254				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.89.

b. Computed only for a 2x2 table

Age & English proficiency

Completers were 6 years younger (31) than non completers (38) ($p<0.046$)

Group Statistics

	Gr.EligCert	N	Mean	Std. Deviation	Std. Error Mean
tpeq.age	Not eligable (not completed)	61	37.98	14.344	1.837
	Eligable (completed)	28	31.79	11.057	2.090
tpeq.engprof	Not eligable (not completed)	64	4.11	1.056	0.132
	Eligable (completed)	28	4.25	0.585	0.111

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
tpeq.age	Equal variances assumed	2.658	0.107	2.025	87	0.046	6.198	3.061	0.113	12.282
	Equal variances not assumed			2.228	66.872	0.029	6.198	2.782	0.645	11.751
tpeq.engprof	Equal variances assumed	5.919	0.017	-0.660	90	0.511	-0.141	0.213	-0.564	0.282
	Equal variances not assumed			-0.817	84.876	0.416	-0.141	0.172	-0.483	0.202

Gender

Crosstab

			tpeq.gender		
			Male	Female	Total
Gr.EligCert	Not eligable (not completed)	Count	42	20	62
		% within Gr.EligCert	67.7%	32.3%	100.0%
	Eligable (completed)	Count	15	13	28
		% within Gr.EligCert	53.6%	46.4%	100.0%
Total		Count	57	33	90
		% within Gr.EligCert	63.3%	36.7%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.668[a]	1	0.197		
Continuity Correction[b]	1.114	1	0.291		
Likelihood Ratio	1.644	1	0.200		
Fisher's Exact Test				0.240	0.146
Linear-by-Linear Association	1.649	1	0.199		
N of Valid Cases	90				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.27.

b. Computed only for a 2x2 table

Highest obtained degree

Crosstab

			tpeq.degree			
			Professional degree or lower	Bachelor's degree	Master's degree or higher	Total
Gr.EligCert	Not eligable (not completed)	Count	13	19	26	58
		% within Gr.EligCert	22.4%	32.8%	44.8%	100.0%
	Eligable (completed)	Count	5	14	9	28
		% within Gr.EligCert	17.9%	50.0%	32.1%	100.0%
Total		Count	18	33	35	86
		% within Gr.EligCert	20.9%	38.4%	40.7%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	2.397[a]	2	0.302
Likelihood Ratio	2.372	2	0.305
Linear-by-Linear Association	0.214	1	0.644
N of Valid Cases	86		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.86.

Current occupation

Crosstab

			tpeq.occupt							
			Student	Recently graduated	Working professional	Looking for a job	Parent / caregiver	Retired	Other	Total
Gr.EligCert	Not eligable (not completed)	Count	8	1	42	5	1	4	0	61
		% within Gr.EligCert	13.1%	1.6%	68.9%	8.2%	1.6%	6.6%	0.0%	100.0%
	Eligable (completed)	Count	4	0	20	2	0	0	2	28
		% within Gr.EligCert	14.3%	0.0%	71.4%	7.1%	0.0%	0.0%	7.1%	100.0%
Total		Count	12	1	62	7	1	4	2	89
		% within Gr.EligCert	13.5%	1.1%	69.7%	7.9%	1.1%	4.5%	2.2%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	7.176[a]	6	0.305
Likelihood Ratio	9.224	6	0.161
Linear-by-Linear Association	0.020	1	0.888
N of Valid Cases	89		

a. 11 cells (78.6%) have expected count less than 5. The minimum expected count is 0.31.

Continent of origin

Crosstab

			tpeq.continent						
			Europe	North America	South America	Africa	Asia	Australia&Oceania	Total
Gr.EligCert	Not eligable (not completed)	Count	21	10	7	6	16	2	62
		% within Gr.EligCert	33.9%	16.1%	11.3%	9.7%	25.8%	3.2%	100.0%
	Eligable (completed)	Count	11	4	2	2	9	0	28
		% within Gr.EligCert	39.3%	14.3%	7.1%	7.1%	32.1%	0.0%	100.0%
Total		Count	32	14	9	8	25	2	90
		% within Gr.EligCert	35.6%	15.6%	10.0%	8.9%	27.8%	2.2%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.854[a]	5	0.869
Likelihood Ratio	2.460	5	0.783
Linear-by-Linear Association	0.047	1	0.828
N of Valid Cases	90		

a. 5 cells (41.7%) have expected count less than 5. The minimum expected count is 0.62.

Enrolment intention

Crosstab

			tpeq.intention		
			I want to complete this course	I do not intend to complete this course	Total
Gr.EligCert	Not eligable (not completed)	Count	53	11	64
		% within Gr.EligCert	82.8%	17.2%	100.0%
	Eligable (completed)	Count	27	1	28
		% within Gr.EligCert	96.4%	3.6%	100.0%
Total		Count	80	12	92
		% within Gr.EligCert	87.0%	13.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.184[a]	1	0.074		
Continuity Correction[b]	2.097	1	0.148		
Likelihood Ratio	3.886	1	0.049		
Fisher's Exact Test				0.098	0.066
Linear-by-Linear Association	3.149	1	0.076		
N of Valid Cases	92				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.65.

b. Computed only for a 2x2 table

Enrolment motivation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gr.EligCert tpeq.enrolmot	*91	67.9%	43	32.1%	134	100.0%

Gr.EligCert * tpeq.enrolmot Crosstabulation

			tpeq.enrolmot		
			To use in daily work activities or improve career prospects	Interested in subject and want to learn more	Total
Gr.EligCert	Not eligable (not completed)	Count	38	27	65
		% within tpeq.enrolmot	69.1%	75.0%	71.4%
	Eligable (completed)	Count	17	9	26
		% within tpeq.enrolmot	30.9%	25.0%	28.6%
Total		Count	55	36	91
		% within tpeq.enrolmot	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	0.372[a]	1	0.542		
Continuity Correction[b]	0.139	1	0.709		
Likelihood Ratio	0.376	1	0.540		
Fisher's Exact Test				0.638	0.357
Linear-by-Linear Association	0.368	1	0.544		
N of Valid Cases	91				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.29.

b. Computed only for a 2x2 table

Assignments made

Paying students

Significant ($p < 0.017$) difference between grade obtained by students who paid and students who did not pay.

Group Statistics

	EnrollTrack	N	Mean	Std. Deviation	Std. Error Mean
Gr.MadeAssign	Not paying students	237	2.73	1.821	0.118
	Paying students	17	3.82	1.704	0.413

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Gr.MadeAssign	Equal variances assumed	1.719	0.191	-2.392	252	0.017	-1.089	0.455	-1.986	-0.193
	Equal variances not assumed			-2.534	8	18.71 0.020	-1.089	0.430	-1.990	-0.189

Age & English proficiency

		Number of assignments made
tpeq.age	Pearson Correlation	-0.213*
	Sig. (2-tailed)	0.045
	N	89
tpeq.engprof	Pearson Correlation	0.059
	Sig. (2-tailed)	0.576
	N	92

Gender

Group Statistics

	tpeq.gender	N	Mean	Std. Deviation	Std. Error Mean
Number of assignments made	Male	57	2.26	1.620	0.215
	Female	33	3.00	1.854	0.323

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number of assignments made	Equal variances assumed	3.184	0.078	-1.971	88	0.052	-0.737	0.374	-1.480	0.006
	Equal variances not assumed			-1.9012	59.87	0.062	-0.737	0.388	-1.512	0.039

Highest obtained degree

Descriptives

Number of assignments made

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Professional degree or lower	18	2.39	1.685	0.397	1.55	3.23	1	5
Bachelor's degree	33	2.85	1.856	0.323	2.19	3.51	1	5
Master's degree or higher	35	2.49	1.687	0.285	1.91	3.07	1	5
Total	86	2.60	1.744	0.188	2.23	2.98	1	5

ANOVA

Number of assignments made

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.295	2	1.648	0.536	0.587
Within Groups	255.263	83	3.075		
Total	258.558	85			

Current occupation

Descriptives

Number of assignments made

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Student	12	2.58	1.730	0.499	1.48	3.68	1	5
Recently graduated	1	1.00	.	.	.	.	1	1
Working professional	62	2.66	1.764	0.224	2.21	3.11	1	5
Looking for a job	7	1.86	1.574	0.595	0.40	3.31	1	5
Parent / care giver	1	3.00	.	.	.	.	3	3
Retired	4	1.00	0.000	0.000	1.00	1.00	1	1
Other	2	5.00	0.000	0.000	5.00	5.00	5	5
Total	89	2.55	1.739	0.184	2.18	2.92	1	5

ANOVA

Number of assignments made

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28.362	6	4.727	1.631	0.149
Within Groups	237.661	82	2.898		
Total	266.022	88			

Continent of origin

Descriptives

Number of assignments made

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Europe	32	2.66	1.753	0.310	2.02	3.29	1	5
North America	14	2.43	1.742	0.465	1.42	3.43	1	5
South America	9	2.00	1.732	0.577	0.67	3.33	1	5
Africa	8	2.50	1.773	0.627	1.02	3.98	1	5
Asia	25	2.68	1.842	0.368	1.92	3.44	1	5
Australia&Oceania	2	2.00	1.414	1.000	-10.71	14.71	1	3
Total	90	2.53	1.737	0.183	2.17	2.90	1	5

ANOVA

Number of assignments made

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.313	5	0.863	0.274	0.926
Within Groups	264.087	84	3.144		
Total	268.400	89			

Enrolment intention

Group Statistics

	tpeq.intention	N	Mean	Std. Deviation	Std. Error Mean
Number of assignments made	I want to complete this course	80	2.66	1.764	0.197
	I do not intend to complete this course	12	1.33	0.888	0.256

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number of assignments made	Equal variances assumed	26.084	0.000	2.553	90	0.012	1.329	0.521	0.295	2.363
	Equal variances not assumed			4.110	26.599	0.000	1.329	0.323	0.665	1.993

Enrolment motivation

Group Statistics

	tpeq.enrolmot	N	Mean	Std. Deviation	Std. Error Mean
Number of assignments made	To use in daily work activities or improve career prospects	55	2.67	1.711	0.231
	Interested in subject and want to learn more	36	2.08	1.645	0.274

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number of assignments made	Equal variances assumed	0.672	0.414	1.631	89	0.106	0.589	0.361	-0.129	1.307
	Equal variances not assumed			1.645	77.064	0.104	0.589	0.358	-0.124	1.303

Grade

Paying students

Group Statistics

	EnrollTrack	N	Mean	Std. Deviation	Std. Error Mean
Grade	Not paying students	237	37.77	30.682	1.993
	Paying students	17	58.59	28.498	6.912

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Grade	Equal variances assumed	0.838	0.361	-2.715	252	0.007	-20.820	7.670	-35.926	-5.715
	Equal variances not assumed			-2.894	18.762	0.009	-20.820	7.193	-35.889	-5.752

Age & English proficiency

		Grade
tpeq.age	Pearson Correlation	-0.226*
	Sig. (2-tailed)	0.033
	N	89
tpeq.engprof	Pearson Correlation	0.082
	Sig. (2-tailed)	0.438
	N	92

Gender

Group Statistics

	tpeq.gender	N	Mean	Std. Deviation	Std. Error Mean
Grade	Male	57	31.44	25.884	3.428
	Female	33	45.67	33.209	5.781

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Grade	Equal variances assumed	6.299	0.014	-2.261	88	0.026	-14.228	6.292	-26.732	-1.724
	Equal variances not assumed			-2.117	54.607	0.039	-14.228	6.721	-27.700	-0.756

Highest obtained degree

Descriptives

Grade

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Professional degree or lower	18	31.44	26.066	6.144	18.48	44.41	0	78
Bachelor's degree	33	41.09	30.983	5.393	30.10	52.08	3	95
Master's degree or higher	35	37.74	30.455	5.148	27.28	48.20	5	100
Total	86	37.71	29.682	3.201	31.35	44.07	0	100

ANOVA

Grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1083.875	2	541.938	0.609	0.546
Within Groups	73803.857	83	889.203		
Total	74887.733	85			

Current occupation

Descriptives

Grade

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Student	12	38.17	30.036	8.671	19.08	57.25	0	88
Recently graduated	1	20.00	.	.	.	.	20	20
Working professional	62	37.66	28.794	3.657	30.35	44.97	3	98
Looking for a job	7	30.57	36.032	13.619	-2.75	63.90	0	100
Parent / care giver	1	40.00	.	.	.	.	40	40
Retired	4	9.75	2.363	1.181	5.99	13.51	8	13
Other	2	86.50	4.950	3.500	42.03	130.97	83	90
Total	89	36.84	29.548	3.132	30.62	43.07	0	100

ANOVA

Grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	8499.280	6	1416.547	1.700	0.131
Within Groups	68330.518	82	833.299		
Total	76829.798	88			

Continent of origin

Descriptives

Grade

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Europe	32	39.00	27.715	4.899	29.01	48.99	5	88
North America	14	31.21	27.708	7.405	15.22	47.21	0	78
South America	9	29.67	23.927	7.976	11.27	48.06	10	75
Africa	8	36.13	35.349	12.498	6.57	65.68	5	100
Asia	25	39.92	34.535	6.907	25.66	54.18	0	98
Australia&Oceania	2	26.50	12.021	8.500	-81.50	134.50	18	35
Total	90	36.58	29.489	3.108	30.40	42.75	0	100

ANOVA

Grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1504.383	5	300.877	0.333	0.892
Within Groups	75887.572	84	903.423		
Total	77391.956	89			

Enrolment intention

Group Statistics

	tpeq.intention	N	Mean	Std. Deviation	Std. Error Mean
Grade	I want to complete this course	80	38.88	29.575	3.307
	I do not intend to complete this course	12	16.50	18.701	5.399

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Grade	Equal variances assumed	15.494	0.000	2.539	90	0.013	22.375	8.813	4.866	39.884
	Equal variances not assumed			3.534	20.402	0.002	22.375	6.331	9.186	35.564

Enrolment motivation

Group Statistics

	tpeq.enrolmot	N	Mean	Std. Deviation	Std. Error Mean
Grade	To use in daily work activities or improve career prospects	55	38.02	29.592	3.990
	Interested in subject and want to learn more	36	30.53	27.343	4.557

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Grade	Equal variances assumed	1.071	0.304	1.216	89	0.227	7.490	6.159	-4.747	19.728
	Equal variances not assumed			1.237	79.103	0.220	7.490	6.057	-4.566	19.547