Old school or new school?

Exploring the effect of gamification elements within an Augmented Reality application on students’ motivation when performing an assembly task

Author:	Egbert Roelofsen
Student Number:	S1886290
Supervisors:	Dr. J. Karreman & Dr. S. Janssen
Study:	Master Communication Studies, Technology & Communication
Date:	September, 2020
Place:	University of Twente, Enschede

Abstract

Background and purpose: The effects of gamification elements within an augmented reality application on students' motivation to perform an assembly task have been researched. Augmented Reality and gamification have been researched before, but separately. It is unknown if the motivational properties of Augmented Reality and gamification become stronger when they are combined. This combination is used within an experimental lab setting where indivuals need to complete an assembly task. Self-Determination Theory is used as the framework to define and observe motivation. In addition, it was also examined what possible effect past experience(s) with Augmented Reality and an assembly task can have on the participants' performance.
Method: Three manipulations existed next to a plain condition that served as a baseline. Eventually, 106 students were randomly assigned to either the plain, leaderboard, badges, or levels condition. As Self-Determination Theory was used as a framework, the Basic Psychological Needs Scale to measure autonomy and competence satisfaction, and the Situational Motivational Scale to determine the locus of motivation were used.
Results: The only found main effect was between gamification and the level of uncertainty. The difference was due to the leaderboard condition having a significantly lower level of uncertainty in comparison to all other conditions. It was explored if competence satisfaction mediated the main effect, which was concluded not to be the case. Furthermore, it was found that past experience with Augmented Reality and past experience with similar tasks had no significant effect on the performance of the participants.
Conclusion: The present study explored how the motivation of students is influenced when gamification and Augmented Reality are combined. It was found that only the leaderboard condition influenced the level of uncertainty of the participants. All other gamification elements and none of the past experience variables affected the motivation or performance of the participants. As this was a first study of its kind, multiple suggestions for future research are discussed.

Table of Contents

1. Introduction

Teaching is an ancient part of human society. Schools and schooling systems have similarly existed for hundreds of years. They primarily exist to transfer the knowledge of the teacher to the students. This brings that students need to be motivated because a lack of motivation can lead to absence (Meijers, 2006) or dropping out entirely (Dresel & Grassinger, 2013). And when a student is not present, the teacher cannot transfer their knowledge. According to the Dutch Ministry of education, culture and science (2020), around 18% of the scientific-educational level students within the Netherlands leave university without a degree. When looking at the general higher educational level within the Netherlands, this metric increases to 38%. Nonetheless, between one and two students drop out of the higher educational levels in the Netherlands on average. New technologies hold the promise of improving teaching methods, like Augmented Reality (AR) or making use of gamification elements like points, badges, and levels. However, to ensure the best possible results, a student needs to be motivated (Hornstra et al., 2016; Vansteenkiste et al., 2007) and new technologies like AR can support and enhance motivation (Radu, 2014) and, thus, possibly decrease the chance of dropping out.

AR offers the ability to support real-world experience with virtual additions (Y.-C. Chen, 2006). It enables the user to see their surroundings and receive extra information needed for their situation. AR has been researched for its ability to enhance motivation and being able to increase the users' performance (Akçayır & Akçayır, 2017; Altinpulluk, 2019). At the same time, gamification, using game-like mechanisms in contexts that are not games, has recently become the focus area of multiple studies to determine if it can improve motivation (Buckley & Doyle, 2016; Dahlstrøm, 2017; Erbas & Demirer, 2019; Mekler et al., 2017; Tröndle, 2016). Both AR and gamification have been reported to be able to increase motivation, but it is unknown if they can reinforce each other's effects. As both AR and gamification are being researched within the field of education, they are bound to intersect. Therefore it is important to observe how AR and gamification interact. When AR and gamification are able to increase the motivation of students, keep them more engaged and decrease the possibility of dropping out, it would mean that more students will finish their study, are able to acquire a higher degree of well-being, and experience a more stable life (Witte et al., 2014).

This research will, thus, focus on gamification strategies applied in an AR application which will be used by university students. By making use of an assembly task the students need to complete in an experimental setting, it is expected that depending on the gamification element, their motivation will increase or decrease. This brings us to the central question of this research:

"To what extent do gamification strategies within an AR application influence the motivation for executing an assembly task of university students?"

To be able to answer the central question, it is necessary first to define motivation through Self-Determination theory. Secondly, Augmented Reality will be defined. Thirdly, gamification will be looked at, and the elements that will be used in the present study are introduced. Lastly, the variable past experience will be discussed.

2. Theoretical Framework

The theoretical framework will dive into the subjects of motivation, Augmented Reality, Gamification, and the effect past experience can have on new situations. Firstly, motivation is being defined and looked at by making use of Self-Determination Theory. Secondly, Augmented Reality will be discussed and explained. Thirdly, gamification is defined and three gamification elements that are being used in the present study are being discussed. Lastly, past experience is being looked at as it can influence how participants respond to new situations.

2.1. Self-Determination Theory

Self-Determination Theory (SDT) is a macro theory of human motivation. It states that every person has a natural focus on personal growth and learning; being naturally intrinsically motivated to learn (Ryan & Deci, 2000). Vansteenkiste et al. (2007) make a comparison with how toddlers have a natural tendency or drive to discover their surroundings and to learn. This is the purest form of intrinsic motivation because the toddler wants to do it solely for itself. This simultaneously addresses the possibility that the toddler could act in a certain way only because a parent would want them to, which would be an example of a toddler being extrinsically motivated. Besides making the differentiation of intrinsic and extrinsic motivation, Deci and Ryan went a step further. They defined multiple levels of motivation in one of their micro theories called Organismic Integration Theory (OIT) (2000).

2.1.1 Organismic Integration Theory

This theory begins with making the distinction between intrinsic motivation, extrinsic motivation, and a third type where no motivation exists, called amotivation. Then the theory takes a closer look at external motivation by specifying four different types. These distinctions have been made visual in Figure 1 (Deci & Ryan, 2000, p. 237). The different types do not state that a person is highly or lowly motivated. It merely describes the locus of said motivation, also considered the 'quality' of the motivation (Vansteenkiste et al., 2007).

Figure 1
Motivation According to Organismic Integration Theory as Part of Self-Determination Theory

Behaviour	Nonself-determined					Self-determined
Type of motivation	Amotivation	Extrinsic motivation				Intrinsic motivation
Type of regulation	Non-regulation	External regulation	Introjected regulation	Identified regulation	Integrated regulation	Intrinsic regulation
Locus of causality	Impersonal	External	Somewhat external	Somewhat internal	Internal	Internal

Note: Reprinted from The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behaviour, Deci, E. & Ryan, R., 2000, *Psychological Inquiry, 11*(11), p. 237

- External regulation; this is the pure form of extrinsic motivation. Students that only attend university to gain prestige or that are being pressured by their surroundings.
- Introjected regulation; in this case, a person experiences internal pressure which drives their behaviour (e.g., feeling guilty or anxious). A student that feels that he or she needs to be able to finish their course.
- Identified regulation; here, the student might not want to write their essay, but they need to pass the course to be able to finish their school, which is their primary goal. This results in writing the essay while they might not be internally motivated to execute the task itself.

- Integrated regulation; this is the next step of identified regulation. Whenever an activity turns into a goal itself and other activities need to take place to achieve said goal, the person integrates the activity and makes it part of their motivation. A student can have the goal to know everything from a specific niche within their domain. When following a boring course is needed to fulfill that goal, following the course belongs to identified regulation. When the student needs to partake in activities within the course, the identified motivation for the course becomes integrated. It enables the student to regulate the new activities making the student identified motivated for them.
- Intrinsic motivation. In this case, a student will partake in a course because they want to. They are not motivated because of wanting to achieve an external goal, but because they want to participate, engage, and collect as much information or knowledge as possible.

Eventually, Ryan and Deci became interested in what determines where someone falls in the motivation continuum (2000). They discovered that humans have three basic psychological needs that are not created but are innate. They described these needs and their interplay in the next micro theory known as the Basic Psychological Needs Theory.

2.1.2 Basic Psychological Needs Theory

The Basic Psychological Needs Theory (BPNT) is a micro-theory within SDT that describes the three basic innate needs, known as autonomy, competence, and relatedness. Autonomy describes the need for being the source of what is happening in one's life. However, Deci & Vansteenkiste (2004) add that this does not mean that someone wants to be completely self-reliant and independent. Instead, it is about acting out of one's own volition and choice. Competence is about feeling competent in what someone does, which applies to both personal and professional life (Deci & Ryan, 2002). Feeling the need to connect with other people and belong to those people and their community is what relatedness is about (Deci & Ryan, 2002; Sun et al., 2017). In essence, this need is about having a connection and being accepted.

Autonomy, competence, and relatedness satisfaction have shown to result in better well-being, better dealing with stress and better performance (Kühne, 2019; Núñez & León, 2015; Vansteenkiste et al., 2010). Within education, need satisfaction has been reported to increase engagement, performance, quality of learning (Núñez & León, 2015), school satisfaction, persistence and lower drop-out rates (Badri et al., 2014; Guay et al., 2008; Ratelle et al., 2007). Need satisfaction has also been linked with improving the well-being of students (e.g., Vansteenkiste & Ryan, 2013). The opposite, called need thwarting, has shown to have multiple down-sides, such as diminishing well-being, being more at risk for burnouts, and more affected by stress (Vansteenkiste et al., 2010). Within education, it has been associated with lower engagement, disengagement, having a more controlled form of motivation, increased drop-out rate, and increased absenteeism (Ratelle et al., 2007; Soenens et al., 2012). Thus, being able to satisfy the basic needs and limit the thwarting of said needs is very important to ensure motivated, well-performing, engaged students that attend classes and have a higher level of well-being. Satisfying the basic needs will result in the student being intrinsically motivated. Deci and Ryan moved on to their next micro theory called the Cognitive Evaluation Theory that tries to determine which factors increase or diminish the satisfaction with those basic needs and, thus, intrinsic motivation.

2.1.3 Cognitive Evaluation Theory

Another micro-theory of SDT is the Cognitive Evaluation Theory (CET), which looks at what kind of factors contribute to or diminish intrinsic motivation (Vansteenkiste et al., 2010). Being able to determine if something will contribute to or diminish intrinsic motivation, helps in formulating accurate hypotheses. First and foremost, CET states that whenever all three basic psychological needs are satisfied, a person will become intrinsically motivated (Deci & Ryan, 2000). At the same time, however, autonomy and competence need satisfaction are seen as the main contributors to fostering intrinsic motivation (Peters et al., 2018; Ryan et al., 2006; Vansteenkiste & Ryan, 2013). Furthermore, Intrinsic motivation can be diminished by extrinsic rewards and could even be crowded out when extrinsic motivators are (too) strong, and this effect is known as the crowding-out effect (Deci, 1972; Deci & Ryan, 1985; Lepper & Greene, 1978).

CET does state that the way someone perceives the extrinsic motivator can mediate the crowding-out effect (Deci & Ryan, 1985, 2000). This boils down to if the motivator is perceived as controlling or informing (Vansteenkiste et al., 2010). Informative elements foster intrinsic motivation, and controlling elements generally result in extrinsic motivation. Elements are perceived as controlling when, for example, they are something tangible, a reward, and are expected (Vansteenkiste et al., 2010). Receiving a bonus after an employee performed well is a form of a controlling element, which functions as an extrinsic motivator. Whenever an element is not expected and not task-dependent – i.e. you do not need to complete the task in order to receive or gain something else – and the element gives feedback about the performance of the person, it is an informative element, which functions as an intrinsic motivator (Vansteenkiste et al., 2010). Receiving feedback about a completed task without it resulting in a reward or punishment would be an example of an informative element which functions as an intrinsic motivator.

Where CET gives the possibility to know what kind of an effect an element might have, the meta-analysis of Cerasoli et al. (2014) directly shows which motivators will have what kind of effect, depending on the task they are being used on. They found that whenever a task is concerned with quality, intrinsic motivators, also known as informative elements, improved performance the most (Cerasoli et al., 2014). These types of tasks are often associated with being more personally invested and having less need for externalized motivation (Deci & Ryan, 2000; Ryan & Deci, 2000). Extrinsic motivators, also known as controlling elements, were found to be best used to increase performance when a task was concerned with quantity. Quantity tasks are typically repetitive and have lower complexity, which results in lower personal involvement (Gilliland & Landis, 1992). As stated in the introduction, the present study uses an assembly task, which means that the quality of the execution of the task is essential. Besides, the effectiveness of the task is only influenced by the quality of the execution. This gives that the assembly task is concerned with quality and, thus, intrinsic motivators are expected to increase performance whereas extrinsic motivators are expected to decrease performance.

2.2 Augmented Reality

Whenever technology is being used to add information which has been generated by a computer to the real world, we talk about Augmented Reality (AR) (Bottani & Vignali, 2019; Krevelen & Poelman, 2010). This addition of information creates the distinction with Virtual Reality, i.e., having a simulation of a fictitious or real-life environment (Khan et al., 2011). AR can primarily be used in two differing ways: via Head-Mounted Displays and via handheld devices. Head-Mounted Displays are units that people wear on their head, where they look through special lenses on which the digital information is being projected. These displays are expensive and more challenging to develop an application for. Handheld devices, like mobile phones and tablets, are the easiest way to use AR. Altinpullik (2019) and Akçayır & Akçayır (2017) also showed that tablets and smartphones are being used the most in current AR research. These mobile devices frequently either make use of a marker or are markerless (Pellas et al., 2018; Saltan & Arslan, 2017). Garzón et al. (2019) add that AR within education is distinctively used for the creation of applications that aim to enhance learning and teaching. Sommerauer and Müller (2018) showed that within education, the focus lies on the task that is being performed and that people learn via their experiences, which makes AR perfect for use within educational contexts.

AR is a good contender to increase the need for autonomy of the user. Garzón et al. (2019) showed that this effect had been mentioned in 26% of all observed papers in their literature review.

This makes sense, as AR enables users to walk around and exercise their autonomy to determine how they can use AR in a way that suits them best. Paper instructions, for example, are static and need to convey all the information through text or 2D visuals, whereas AR adds 3D visuals that can be seen from every side. The 3D information can also assist the user in grasping the meaning of the shown information (Yoon et al., 2012). If an instructional step would be vague in the paper instruction, the AR instruction enables the possibility to change the perspective of the user until the instruction would become apparent. Buchner & Zumbach (2018) compared the level of perceived choice, which is seen as a measure for autonomy within SDT, between a group that worked with AR and a group that did not. The difference in perceived choice was significant, which underlines the idea that AR satisfies the need for autonomy. When we take all the preceding information, AR can be seen as an exciting tool for educational purposes.

2.3 Gamification

Gamification has often been defined as "… the use of game design elements in non-game contexts." (Deterding et al., 2011, p. 10) and has been researched quite frequently in recent years (Barata et al., 2017; Dahlstrøm, 2017; Hung et al., 2017). The area that is mainly popular within this body of research is the field of education, see, for example, the literature review of Hamari, Koivisto, and Sarsa (2014) and Seaborn and Fels (2015). The found effects, however, have been mixed (Dahlstrøm, 2017; Hamari et al., 2014; Seaborn & Fels, 2015). Effects are sometimes positive, sometimes negative, and sometimes neutral. Seaborn and Fels (2015) saw that the effectiveness of gamification is dependent on the educational area it is being used in, how acquainted the users are with playing games, and their age. Personal characteristics and preferences are other important contributors to explain the differing results, according to Hamari et al. (2014). As of now, it is unknown what characteristics are responsible for the successful implementation of gamification. Therefore, it is an area that needs to be continuously researched in order to unearth when gamification elements are beneficial, and when they are detrimental. Dahlstrom (2017) does conclude that gamification brings positive effects to some degree, albeit being very dependent on the domain it is used in and the characteristics of the people that are exposed to gamification.

The literature review of Seaborn and Fels (2015) shows at the same time that the possible reasons for the mixed results have not been clearly studied. An explanation Seaborn and Fels (2015) and Hamari, Koivisto, and Sarsa (2014) give is that certain promising elements work in one given domain, but not in another. This gives the notion that the context of the research is essential, however, both Seaborn and Fels (2015) and Hamari et al. (2014) did not mention wat kind of context has wat kind of an effect. This makes it necessary to research the application of gamification in multiple contexts to discover when it is effective and when it has no effect at al.

Furthermore, it is important to know what kind of elements are generally researched. Hamari et al. (2014) give an overview, which shows that points, leaderboards, badges, and levels are often researched. At the same time, Seaborn and Fels (2015) add that current studies tend to combine multiple gamification elements, which make it challenging to know which effect(s) the various elements have. The present study, therefore, will focus individually on leaderboards, badges, and levels.

Another point of critique given by Seaborn and Fels (2015) is the lack of validated theoretical foundation and the lack of studies that use validated scales. There has also been little research into what effect which gamification element has. Mekler et al. (2017) have made a start by studying if gamification elements like points, levels, and a leaderboard increased performance and intrinsic motivation within an image annotation task, a task mainly concerned with quantity. They found that the gamified group all produced more annotations, which were not of a higher quality than the control group. Another interesting conclusion was that intrinsic motivation was not affected by the gamification elements (Mekler et al., 2017). This matches the conclusions of Cerasoli et al. (2014) that states that tasks concerned with quantity can be enhanced in terms of performance by making primarily use of extrinsic motivators.

2.3.1 Types and effects of Gamification

Commonly used gamification elements are leaderboard, badges, and levels (Majuri et al., 2018). This study will make use of leaderboard, badges and levels as the gamification techniques that are being observed.

Leaderboard. A leaderboard works by setting a goal for the participant to attain and providing feedback in terms of being able to obtain points representing the individuals' level of competence (Jung et al., 2010). The leaderboard in the present study works with the same principle because the participants' actions and time to completion gives a final score that represents their level of competence. Jung et al. (2010) and Mekler et al. (2017) both found that a leaderboard resulted in a significant boost in performance, and both experimental tasks were also concerned with quantity. As this study uses a task that is concerned with quality, it is unclear if the same effect can be expected. It is expected that the leaderboard is seen to be controlling in nature, "I want to be the best", as opposed to informative, "I want to know how well I participated in regard to others" (Vansteenkiste et al., 2010). It is also expected that the task in the present study will be seen as challenging due to it being a specialized assembly task. Wang et al. (2015) found that controlling feedback in a challenging condition resulted in a significantly lower level of perceived competence and, thus, it is expected that the element will be competence need thwarting. This brings us to the following hypotheses:

H1a: The use of a leaderboard is negatively related to students' competence satisfaction in completing the task in comparison with the no-manipulations condition.
H1b: The use of a leaderboard will negatively influence autonomous motivation in comparison with the no-manipulations condition.
H1c: The use of a leaderboard will positively influence controlled motivation in comparison with the no-manipulations condition.
H1d: The use of a leaderboard is negatively related to students' performance in comparison with the no-manipulations condition.
H1e: The effect of a leaderboard on motivation and performance is mediated by competence satisfaction.

Badges. Badges can be seen as a form of praise due to achievement generally giving "a positive evaluation of one's performance" (Hakulinen et al., 2015, p. 19), where Morris et al. (2013) recognize praise as an essential feature that makes games motivating. Furthermore, badges have been shown to be able to improve performance quantity (Denny, 2013) and the motivation of students (McDaniel et al., 2012). It has to be noted that the benefits regarding motivation were reported, but not statistically verified. Additionally, Hanus and Fox (2015) found that badges lowered motivational levels. However, they combined badges with a leaderboard, made the badges visible on the leaderboard, and the participants needed to fill in forms to receive the badges. This made the badges tangible and an extrinsic incentive and, thus, controlling and intrinsic motivation diminishing according to CET (Vansteenkiste et al., 2010). Wang et al. (2015) found that informative elements in a challenging condition resulted in a significantly higher level of competence satisfaction as opposed to controlling elements. The assembly task in the present study is seen as challenging. It is, therefore, expected that badges would increase competence satisfaction. This results in the following hypotheses:

H2a: The use of badges is positively related to students' competence satisfaction in completing the task in comparison with the no-manipulations condition.
H2b: The use of badges will positively influence autonomous motivation in comparison with the no-manipulations condition.

H2c: The use of badges will negatively influence the controlled motivation in comparison with the no-manipulations condition.
H2d: The use of badges is positively related to students' performance in comparison with the no-manipulations condition.
H2e: The effect of badges on motivation and performance is mediated by competence satisfaction.

Levels. Levels function by giving the participant goals which they need to achieve (Jung et al., 2010). In the present study, levels are implemented by dividing the assembly task into multiple stages, where the goal of each stage is to complete all the shown steps. Mekler et al. (2017) found that levels did increase performance in an image annotation task but did not increase intrinsic motivation. Following the result of the meta-analyses of Cerasoli et al. (2014), this would mean that levels need to be an extrinsic motivator to be able to explain the gain in performance but no change in intrinsic motivation levels. This is also in line with statements of Hamari et al. (2014), Kumar (2013), and Reeves and Read (2009), which all state that levels are often used to increase externalized motivation (Aparicio et al., 2012). However, two interesting things are happening. First, it is unusual for levels to be an extrinsic motivator because they give direct feedback, which is generally seen as informative, which fits an intrinsic motivator (Vansteenkiste et al., 2010). Secondly, if levels are an extrinsic motivator, they are also controlling as opposed to informative, which means that they should diminish competence satisfaction in a challenging condition (Wang et al., 2015). As the assembly task at hand is seen as challenging, it is expected in the present study that levels will diminish competence satisfaction. This brings us to the following hypotheses:

H3a: The use of levels is negatively related to students' competence satisfaction in completing the task in comparison with the no-manipulations condition.
H3b: The use of levels will negatively influence autonomous motivation in comparison with the no-manipulations condition.
H3c: The use of levels will positively influence controlled motivation in comparison with the no-manipulations condition.
H3d: The use of levels is negatively related to students' performance in comparison with the no-manipulations condition.
H3e: The effect of levels on motivation and performance is mediated by competence satisfaction.

2.4 Experience

Being able to transfer prior experience to new situations has been researched for some time (Gentner, 1983; Gick & Holyoak, 1983), where the mechanism *analogical transfer* is of foremost relevance for the present study (Nokes, 2009). Analogical transfer describes three stages: remembering a past experienced example of the current situation or problem, map the past example unto the current situation or problem, and lastly, drawing a conclusion based on the mapping that fits the current situation or problem (Holyoak et al., 1994).

In terms of prior AR experience, analogical transfer is something that could very well occur. For example, the use of mobile AR in the present study can look similar to an application of mobile AR participants encountered in the past. They would then map how the situation back then looked like, how the technology worked, and what they needed to do to control the technology. Then they apply the mapping to the new situation and conclude on how to use the mobile AR application. If their past experience with AR were awkward, clunky, or challenging, they would be influenced by that experience. Same goes the other way: when their experience was exquisite, easy to use, and understand, they will most likely expect the current mobile AR application to be on par. Any deviation of their expectation can change how they interact with the technology and, thus, how the manipulations might influence them.

Furthermore, it is not surprising that experience with AR or assembly tasks can influence how effectively people can perform in a new situation. However, this is something that Hamari et al.

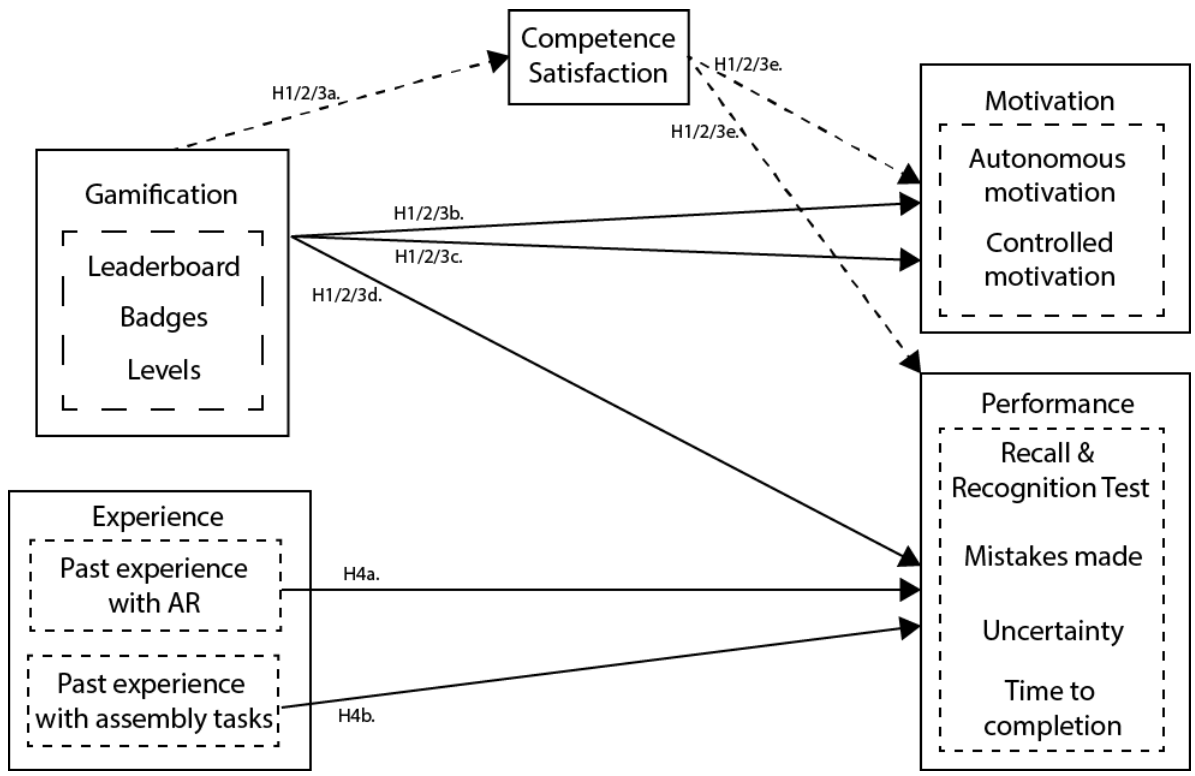
(2014) mentioned in their literature review of AR studies as a methodological limitation in the studies they had reviewed. Moreover, besides Hamari et al. (2014), it has not been mentioned in more recent literature reviews of AR as an influencer of performance (Akçayır & Akçayır, 2017; Garzón et al., 2019; Garzón & Acevedo, 2019).

Taking everything into account, this brings us to the need to control for the possible influence of an individual liking or disliking their previous encounter with AR, and to measure if having previous experience with AR and an assembly task makes them perform better than people with no to little experience. This results in the following two hypotheses, which focus on the possible effect of past experience on performance. To finish this chapter, Figure 2 shows the research model that is being used in the present study.

H4a: Prior experience with Augmented Reality will positively influence the performance of the participants.
H4b: Prior experience with an assembly task will positively influence the performance of the participants.

Figure 2

Research Model of Associations Between Gamification, Competence Satisfaction, Motivation and Performance, and Experience



3. Methodology

3.1 Experimental Design

This study made use of an in-person experiment with the help of a real demo and the Qualtrics survey software. The study consisted of four conditions. The first (control) condition foregoes the use of any gamification elements and served as a baseline of comparison. The second condition contained a leaderboard, the third condition consisted of badges, and the fourth condition used levels. Participants were randomly assigned to each condition. The dependent variables were the locus of motivation and performance and were only measured after the experiment. This makes this research a post-test only randomized between-subjects experiment (Dooley & Vos, 2008).

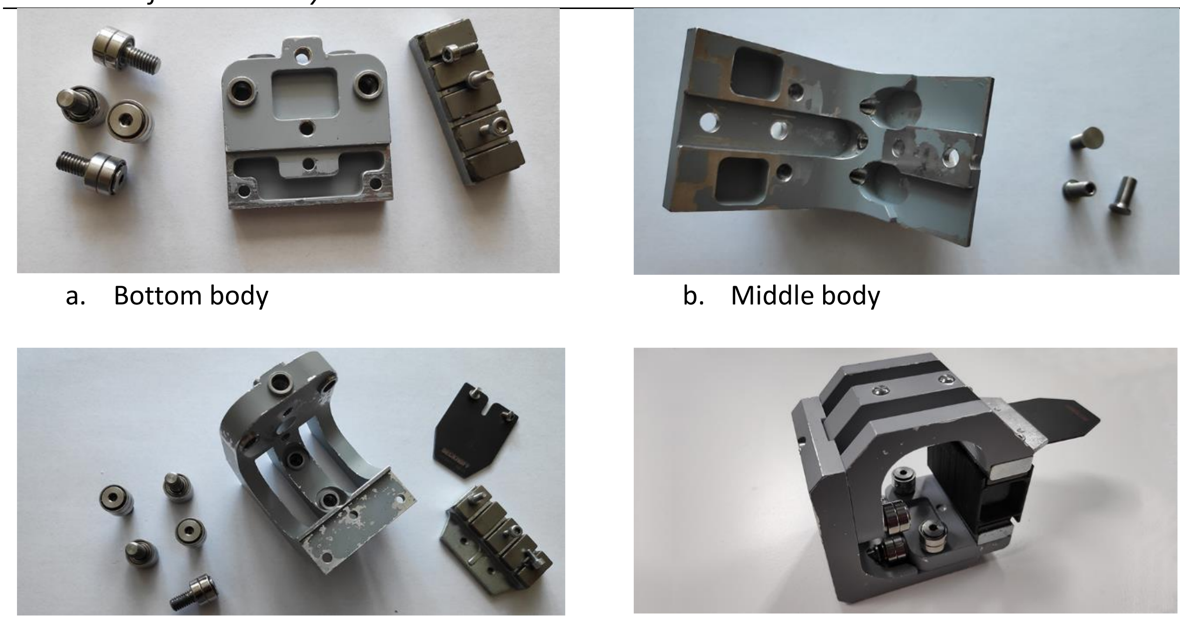
3.2 Stimulus Material Development, Pre-testing, and Manipulation Check

3.2.1 Assembly Task

The piece of equipment the participants needed to assemble has been depicted in figure 3. The assembly task can be divided into four stages: three stages are about preparing every individual metallic body by adding the bolts and the bearings, which are shown in image a, b and c, whereas the last stage is about assembling all three metallic bodies into one which will result in image d. This assembly happens by assembling the bottom and top body to the middle body with two bolts per body, using two lock rings per bolt and a small protector plate per body.

Figure 3

Overview of the Assembly Task Phases



a. Bottom body

b. Middle body

c. Top body

d. Fully assembled

3.2.2 Development of the AR Application

This study made use of an instructional application that had already been developed by a third party. The third party acted as an expert throughout the study when it came to the AR application. They primarily build AR and VR solutions for companies. They have created an AR demo to show the possibilities of AR when meeting with possible clients. That demo has been used for this research. The demo shows how to assemble a piece of equipment that is being used within the food industry, which is shown in Figure 2. Four versions of the application existed due to the differing manipulations. Furthermore, the application is an example of mobile AR, due to its functioning via a tablet. The third party created the three manipulations together with the researcher. The AR application was a marker-based solution. It used an unique identifying image through which the application knew where to display the 3D information.

3.2.3 The Leaderboard Condition

At the start of the task, the participant was shown a screen that informed them about the leaderboard and what would influence their score. During the assembly task, a timer was also shown. When the task was finished, the participant saw a screen that informed them that they were slower than the current top three and, thus, did not receive a position in the leaderboard. This screen was the same for every participant, which results in the same situation for every participant and, thus, mitigating the possible influence of the achieved rank on motivation levels and performance.

To ensure that every participant was aware of this manipulation, a validation question was asked in the questionnaire afterward: “Did the just used AR application contain a leaderboard?”.

3.2.4 The Badges Condition

Based on the difficulty people can experience with assembling components, badges had been implemented to make the user aware of themselves acquiring a new skill. There were two badge types: one type that acknowledges the progression made by the participant, and another that makes the participants aware of mastering a new skill.

Badges were communicated in a textual and visual way. At the start, the respondent was told in a textual way that there are multiple badges present in the game. They were asked if they could find them all. While progressing through the task, participants would receive badges after a predefined number of steps. All badges, including when they are given to the participant, can be found in Figure 4. The manipulation was checked with a simple, straight forward question: “Did the just used AR application contain badges?”.

Figure 4

All Used Badges with their Name and when they are Received

Badges				
Name	Beginner assembler	Master assembler	Beginner tooler	Master tooler
When received	After completing step 4	After the last assembling step	After completing step 8	After completing step 16

3.2.5 The Levels Condition

Due to the used piece of equipment in the present study consisting of three main components, it was deemed easiest to make every component a level on its own. The assembly of all three components into one was the fourth and final level. The number of tasks within each level varied. The first level was determined the easiest by the researcher and the third party. The third party acted as an expert as they have experience with multiple people that have assembled the piece of equipment before. The component with only three tasks was set as the second level because it is somewhat difficult to complete if the participant has no prior experience with using tools. The third level was a copy of the first level, with more steps added to it. Lastly, the fourth level was the last level in which every component needed to be connected.

Levels were communicated in a textual way. At the start of the assembly task, a screen was shown on the mobile AR application that informed the participant that they will start in level one and that the task will be completed after finishing level four. After every level, the participant was shown a screen that indicated that they completed the level and were entering the next level or were finished with the task. To see if the participants were aware of this manipulation, a simple question was included in the questionnaire presented afterward: “Did the just used AR application contain levels?”. The participant could only answer with yes or no.

3.2.6 Pre-testing

A pre-test was held with five participants to ensure that the manipulations and the developed scales were clear to the participants. Every participant was placed in one of the four conditions and progressed through the experiment in the same way real participants would. After they completed the assembly task and filled in the survey, their feedback was collected. It became apparent that the manipulation question was unclear. The questions before the manipulation check inquired if the participant had used AR before and, if applicable, how they liked their previous experience. The pre-testers felt like the manipulation check was referencing their previous experience as opposed to the just used application. The wording was changed from “Did the AR application contain …” to “Did the **just used** AR application contain …”. Furthermore, it was determined to add a visible timer in the leaderboard condition. The badges and levels conditions both had moments where the task was interrupted to inform the participant of either a badge being found or of a level being completed. By contrast, the leaderboard had no elements to remind the participant of the condition; hence a timer was added as a visual reminder.

3.2.7 Manipulation Check

A manipulation check was conducted for every gamified condition to observe if every participant was aware of the condition they were in. Every question was prefixed with “Did the **just used** AR application contain…”. For the levels condition it was “levels?”, for the badges it was “badges?”, and for the leaderboard condition it was “a scoreboard?”. After the first twenty participants, individuals seemed to struggle with the leaderboard manipulation check. After some discussion, it became apparent that the term *scoreboard* did not appear anywhere during the execution of the task. In contrast, the term badge and level were frequently visible in the badges and levels conditions. Those terms also describe what the participants could expect. In the case of the leaderboard condition, it was determined that the measuring of the participants' score was the most distinctive characteristic. Therefore, the manipulation check regarding the leaderboard condition was changed from “a scoreboard?” to “measure your score?”. Table 1 shows the results of the manipulation checks.

Table 1
Results of the Manipulation Checks

Did the just used AR application…		None	Levels	Badges	Leaderboard	Total
Measure your score?	Yes	3	2	1	26	32
	No	23	25	25	1	74
Contain badges?	Yes	3	2	26	5	36
	No	23	25	0	22	70
Contain levels?	Yes	9	26	9	10	54
	No	17	1	17	17	52

3.3 Measurements

3.3.1 Locus of Motivation

Locus of motivation is at the heart of Organismic Integration Theory. One way to determine the type of motivation that can be assigned to a person is via the Situational Motivational Scale (SIMS) (Guay et al., 2000). This scale makes use of four items per subscale, which results in a total of 16 items. This self-report questionnaire measures intrinsic motivation, identified regulation, external regulation, and amotivation. The participants were asked to rate every statement on a 7-point Likert scale ranging from 1 “corresponds not at all” to 7 “corresponds exactly”. The statements are prefaced with the question why the participant is currently engaged in the activity. Examples of the statements are “I do this activity, but I am not sure it is a good thing to pursue it” and “Because I think that this activity is interesting”. The used questions can be found in Appendix 1A.

A different way to test if one of the conditions is perceived as more autonomous or controlling is through breaking the SIMS into two parts by categorizing external regulation and introjection as controlled motivation and the others as autonomous motivation, like Gagné et al. (2010) suggested. This is also in line with the divide of motivation according to the theory, see, for example, Vansteenkiste et al. (2010). The present study also made use of the ability to break the SIMS into two parts per the suggestion of Gagné et al. (2010).

The first performed analysis was a Principal Component Analysis (PCA) followed with an internal consistency check, known as Cronbach’s Alpha. The PCA was run, and an eigenvalue greater than 1 first resulted in five components, as can be seen in appendix 3a. Question 10, a measurement for External Regulation, loaded very negatively (-.788) together with question 11 (.881), a measurement of Amotivation, on a fifth component. It was decided to rerun the PCA without question 10. This resulted in four components, as table 2 shows, and has a Kaiser-Meyer-Olkin measure of KMO = .815, which shows sampling adequacy. Together with Bartlett’s Test of Sphericity $X^2(105) = 783.933, p < .001$ the data used was adequate to run a PCA. This final model resulted in explaining 68.94% of the total variance. Then the internal reliability of every construct was measured, which can be also found in table 2, combined with the eigenvalues per motivational type.

Table 2
Principal Components Analysis of the SIMS

Construct/Items	Loadings	Eigenvalue	% of variance	α
Intrinsic Motivation		5.46	36.42	.93
Because I think that this activity is interesting.	0.81			
Because I think that this activity is pleasant.	0.90			
Because this activity is fun.	0.88			
Because I feel good when doing this activity.	0.81			
Internal Regulation		1.20	7.97	.67
Because I am supposed to do it.	0.84			
Because it is something that I have to do.	0.87			
Because I don't have any choice.	0.53			
Because I feel that I have to do it.	0.83			
External Regulation		2.14	14.25	.82
There may be good reasons to do this activity, but personally I don't see any.	0.68			
I do this activity but I am not sure if it is worth it.	0.64			
I don't know; I don't see what this activity brings me.	0.65			
I do this activity, but I am not sure it is a good thing to pursue it.	0.77			
Amotivation		1.54	10.30	.73
Because I am doing it for my own good.	0.84			
Because I think that this activity is good for me.	0.61			
Because I believe that this activity is important for me.	0.70			

Notes. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Loadings larger than 0.50 were retained.

3.3.2 Need for Competence and Autonomy

As shown in the theoretical framework, competence and autonomy satisfaction are measurements to assess the level of motivation someone has. To be able to assess competence and autonomy need satisfaction, the items of the Basic Psychological Need Satisfaction and Frustration scale (BPNSFS) was used (B. Chen et al., 2015). The items measuring need frustration and relatedness need satisfaction were removed from the scale, as those measurements did not attribute anything relevant to the present study. The wording of the scale items of this trimmed down version of the BPNSFS scale was slightly changed as well. The scale asks the respondent about the experiences they have in their own life. To measure the need satisfaction levels pertaining to the executed task, every statement was prefaced with "During the just performed task…".

Furthermore, the tense of every item has been changed from present to past. This to emphasize that the scale was asking about the experiences regarding the earlier performed task. This trimmed-down version of the BPNSF contained four autonomy related statements and four competence related statements. Examples of the autonomy satisfaction related questions are "I felt a sense of choice and freedom in the task I did" and "I felt that my decisions reflect what I really wanted". Examples of the competence satisfaction questions are "I felt confident that I could do it well" and "I felt competent to achieve my goal". An overview of the used scale can be found in appendix 1B.

The above steps also have been followed to confirm the researched constructs for the Autonomy Satisfaction and Competence Satisfaction scale. The PCA showed two components as expected. Those components explained 60.157% of the total variance. Together with a KMO = .760 and Bartlett's Test of Sphericity $X^2(28) = 246.783, p < .001$ the data was deemed adequate. These results can be found in appendix 3b. The eigenvalues and the Cronbach's Alpha values are displayed in table 3.

Table 3
Eigenvalues and Internal Reliability of the Autonomy and Competence Items of the BPNS Scale

	Eigenvalue	% of variance	Cumulative %	Cronbach's Alpha
Competence Satisfaction	3.19	39.88	39.88	.79
Autonomy Satisfaction	1.64	20.52	69.55	.73

3.3.3 Performance

Objective performance of the participants was measured via the completion time, the number of mistakes made while executing the assembly task, level of uncertainty during the assembly task, and through a recognition test after the task had been finished. Completion time and observing the number of mistakes are common ways to measure performance with AR applications, see the literature review by Sommerauer and Müller for an overview (2018). The completion time was measured via the use of a built-in timer. This meant that the time spent on the 'badge achieved' and the 'level past' screens had to be excluded from being timed. The researcher present noted the number of mistakes made.

Uncertainty was measured as a combination of two factors: replaying a step and going back one step. To be sure of what the participant needs to do, or if they think they missed something in a previous step, they can choose to replay the current step or go back one or more steps. The amount of times these buttons were used has been measured within the app itself. These two values have been combined to create a new variable called uncertainty. It is thought that being able to replay the current step or going back a step helps the participant to feel sure in what they need to do or just did. Therefore, it is assumed that these two measurements can be combined.

The recognition test considered out of 4 questions of 2 types: type 1 was related to the sequence of the steps within the assembly tasks, where type 2 was about what tools should be used in the portrayed step. The type 1 questions asked the participant to put three images in the right order of assembly. Type 2 questions asked the participant which tool should be used out of 3 answer options. The participant was not made aware of their answer being correct or not. An overview of the recognition test can be found in Appendix 1C.

3.3.4 Demographic and Sample Characteristics

Lastly, demographic and sample characteristics were also measured to be able to know the basic background of the used sample. Asked questions were regarding the gender of the participant, their age, what study they were enrolled in, if they had past experience with AR, how they liked their past experience with AR if they had any, and if they had past experience with assembly tasks.

3.4 Procedure

The whole experiment consisted of three stages: (1) invitation, (2) the experiment itself – this stage consists out of four conditions where one of which was randomly assigned to the participant –, and (3) post-experiment.

In the first stage, students were recruited for participation in the experiment using a University of Twente system that awards them with partial course credit for their effort. The potential participants were informed on the application webpage about the premise and that the study would have a duration between 45 and 60 minutes. In the premise, it was stated that the students would be challenged to build something special through the use of provided instructions.

The second stage was the experiment itself. Here the participant was brought into a room where the participant could not yet see the parts they needed to assemble. Depending on the condition the participant was in, the used tablet was prepared with a different AR application. There were four possibilities: (a) AR with no gamification elements, (b) AR with a start screen that informed the participant about a leaderboard and how their completion time, replays of the instructions and mistakes would influence their place on the leaderboard and an end screen that shows their final score, (c) AR with a start screen informing that the participant can achieve badges after fulfilling certain tasks a set amount of times, (d) AR with multiple extra screens that show the participant that they enter one of the four levels. Before the used condition could be determined, the participant was provided with the consent form, which can be found in appendix 2. Then, after consent was given, a randomizer within Qualtrics determined the conditions the respondent would participate in, and the tablet was set up, fitting the assigned condition. Lastly, all the parts the respondent needed to assemble were revealed, the researcher left the room, and the participant could start the assembly task.

In the third stage, the participants had finished the assembly task and moved over to a desktop computer through which they answered multiple questionnaires. The questions first measured the locus of motivation through the use of 16 items, using a 5-point Likert scale. After this, the respondent answered eight questions to determine their autonomy and competence need satisfaction. Following these questions, the participants were asked four questions to test their recognition of the performed task, which work as part of the performance measure of the participant. After these questions, the participants were asked three manipulation checking questions to determine whether the participant noticed the used gamification element. Then the participants answered three questions to determine their previous experiences with AR applications, how they enjoyed those experiences, and how experienced they were with assembly tasks. Lastly, the participant was redirected to a debriefing page that gave the option to read a summary about the study they partook in, what was being measured, and that feedback could be given to the researcher. When the computer part was finished, the researcher collected any feedback the participant had and thanked them for taking part in the study.

3.5 Demographic Data and Sample Characteristics

During the data collection, a total of 123 students signed up for this research. Seventeen participants were excluded. Eight of these were a no-show; six for being unable to complete the assembly task within the maximum allocated time of 60 minutes; two closed the AR application; and one who started answering the survey questions before they were done with the assembly task. The number of valid participants the data-analysis started with was 106. The distribution of all participants per condition is shown in table 8.

3.5.1 Gender Distribution

To see if the distribution of gender is not significantly different per condition, Chi-Square goodness of fit test has been conducted. The first test was significant, $X^2(2) = 50.74, p < .001$. The data became normally distributed after removing one person with a gender other than male or female from the Chi-Square goodness of fit test, $X^2(1) = .47, p = .495$. Table 8 gives an overview of the distribution of gender.

3.5.2 Age Distribution

A one-way between-subject ANOVA was conducted to evaluate if age affected the none, levels, achievements, and scoreboard conditions. There was no significant difference of age between the conditions at the $p < .05$ level $[F(3, 101) = 1.49, p = .223]$. The average and standard deviation values per condition are represented in table 8.

3.5.3 Study Faculty

The distribution of participants per faculty has been examined via Chi-Square goodness of fit test to determine if faculty could be a factor in any found or not found effects. This test indicates that participants are not equally distributed between the BMS and the Other faculty category, $X^2(1) = 28.81, p < .001$. It is possible that the higher amount of Other faculty students in the condition Badges can skew the data. This distribution of participants can be found in table 4.

Table 4

Demographics of the Participants Across all Conditions

		None	Levels	Badges	Leaderboard	
						Total
Count		26	27	26	27	106
In %		24,5%	25,5%	24,5%	25,5%	100%
						Total
Faculty	BMS	24	20	15	21	80
	Other	2	7	11	6	26
						Total
Gender	Males	12	9	12	16	49
	Females	14	18	13	11	56
	Other	0	0	1	0	1
						Total
Age	M	21,3	21,2	22,6	21,3	21,6
	SD	2,5	2,6	3,2	2,4	2,7

BMS = Behavioural Management and Social Sciences, Other = combination of all other faculties

3.5.4 Experience

Level of prior experience with AR, the level of liking of past AR experience(s), and the level of prior experience with an assembly task were measured. To take note of previous experiences with AR, the participants were asked “How experienced would you rate yourself in the use of AR applications” on a Likert scale ranging from 1, shown as “No experience at all (Used it 0 times)”, to 5, shown as “A lot of experience (Used it 10+ times)”. In addition to this, the participants were asked, “How well did you like your **previous experience** with AR?” on a Likert scale ranging from 1, shown as “I liked it a great deal”, to 5, shown as “I disliked it a great deal”. However, when someone answered “No experience at all (Used it 0 times)” on the question about their experience with AR, they would skip the question about their level of liking their past experiences with AR. This resulted in 51 individuals who answered the question regarding their level of liking of their previous experiences with AR. Lastly, participants were asked, “How experienced would you rate yourself in the type of task just performed?” on a Likert scale from 1, shown as “Never assembled anything 1”, to 5, shown as “I perform similar tasks every week 5”. A full overview of all answer possibilities can be found in Appendix 1D.

Multiple one-way ANOVA’s were conducted to make sure that the level of prior experience with AR, the level of liking past AR experience(s), and the level of prior experience with an assembly task were all similar across all conditions. As Table 5 shows, none of the means of the experience variables were significantly different between any of the conditions, which means that all conditions have similar levels of the experience variables and that distributing the participants randomly was successful. Furthermore, even though only 51 participants had previous experience with AR, their experience is evenly distributed among the conditions and, thus, do not require to be controlled for as a covariable.

Table 5

Means, Standard Deviations, and One-Way Analyses of Variance of Past Experience with AR, Level of Liking Past AR Experience, and Past Experience with Assembly Tasks

Measure	None M (SD)	Levels M (SD)	Badges M (SD)	Leaderboard M (SD)	F	*p*
Past experience with AR	1.92 (.94)	1.63 (.84)	1.58 (.99)	1.89 (1.16)	(3, 102)	.472
Level of liking past AR experience	1.88 (.62)	1.83 (.58)	1.56 (.53)	2.14 (.77)	(3, 47)	.209
Past experience with assembly tasks	2.46 (.99)	2.67 (.92)	2.58 (1.03)	2.30 (.87)	(3, 102)	.520

3.6 Statistical Analysis

The analysis of the research data has been done by making use of SPSS version 26. The level of significance that was used was 5% or lower. The analysis was divided into four phases. The first phase was turning the used scales into variables. The goal of the second phase was to see if there are direct effects of the type of manipulation used on competence satisfaction, locus of motivation, and performance. Phase three was used to evaluate if competence satisfaction acted as a mediator for locus of motivation, and performance. When such an analysis had any value, it was tested by making use of the Process macro plugin made by Hayes for SPSS 26 (Hayes, 2017). The fourth phase was used to see if any additional measured measurements possibly influenced any variable.

4. Results

To be able to conduct the statistical analysis, the variables autonomous motivation, controlled motivation, recognition total test score, and uncertainty were made. Autonomous motivation is the average of the scores of the two intrinsic motivational types of the Situational Motivation Scale (SIMS), known as intrinsic motivation, and identified regulation. Controlled motivation is the average of the scores of the two extrinsic motivation types of the SIMS, known as external regulation, and amotivation. Recognition total test score was calculated by taking the average amount of correct answered questions, which results in a value between zero and one. Uncertainty has been calculated by taking the average score of the number of times the participant replayed their current step, and the number of times the participant went to a previous step. To ensure that only variables that are conceptually closely related are being tested, one ANOVA and two Multivariate ANOVA's have been conducted. The one ANOVA focuses on competence satisfaction, the MANOVA's focus on motivation and performance. The mediation effect has been measure by making use of the Process macro plugin of Hayes (Hayes, 2017). Lastly, a MANOVA was conducted evaluate the possible effect of the past experience variables on performance. Table 6 shows the descriptive statistics of the measured dependent variables. The hypotheses this study tested were as follows:

H1a/2a/3a – [manipulation] positively relates to participant competence satisfaction;
H1b/2b/3b – [manipulation] [positively/negatively] influences autonomous motivation;
H1c/2c/3c – [manipulation] [positively/negatively] influences controlled motivation;
H1d/2d/3d – [manipulation] influences the performance of the participant;
H1e/2e/3e – if there is a main effect, then the effect of [manipulation] on motivation and performance is mediated by the level of competence satisfaction of the participant;
H4a and b – past experience [with AR/with an assembly task] influences performance.

Table 6

Descriptive Statistics of all Variables for all Conditions

	None	Levels	Badges	Leaderboard
N	26	27	26	27
	M (SD)	M (SD)	M (SD)	M (SD)
Autonomous motivation[a]	4.30 (.89)	4.57 (1.07)	4.42 (1.02)	4.23 (.99)
Controlled motivation[b]	3.37 (.76)	3.13 (1.04)	2.77 (1.06)	3.13 (.75)
Competence satisfaction[c]	3.92 (.82)	3.92 (.87)	4.13 (.66)	3.79 (.58)
Autonomy satisfaction[d]	2.58 (.57)	2.88 (.88)	2.98 (.87)	2.64 (.89)
Recognition total test score[e]	1.42 (.95)	1.59 (1.01)	1.62 (.90)	1.74 (.81)
Mistakes made[f]	1.46 (2.10)	2.15 (2.32)	1.96 (2.27)	1.96 (2.10)
Uncertainty[g]	5.67 (4.56)	5.09 (3.92)	5.98 (4.85)	2.00 (2.41)
Time to completion[h]	34:21.70 (7:41.48)	35:07.47 (7:56.98)	34:14.22 (7:27.30)	33:40.28 (9:08.56)

Note. M = Mean. SD = Standard Deviation.

[a+b] Autonomous motivation and controlled motivation are a combination of the measured types of motivation through the use of the Situational Motivational Scale, which ranged from 1 (Strongly disagree) to 7 (Strongly agree).

[c+d] Competence satisfaction and autonomy satisfaction have been measured through the use of the Basic Psychological Need Satisfaction Scale, which ranged from 1 (Not true at all) to 5 (Completely true).

[e] Recognition total test score is the total correctly answered question with a minimum of 0 and a maximum of 4.

[f] Mistakes made are the average number of mistakes made during the assembly task.

[g] Uncertainty is the average number of times an individual replayed an instruction or went back an instruction.

[h] Time to completion is the average time individuals in the respective condition needed to complete the assembly task.

4.1 Competence Satisfaction

Hypotheses 1a, 2a, and 3a were analysed by doing an ANOVA analysis. This ANOVA measured the effects of the gamification elements on competence satisfaction. This resulted in a statistically insignificant difference in competence satisfaction based on differing gamification elements (F (3,102) = .93, p = .420).

4.2 Locus of Motivation

Hypotheses 1b and c, 2b and c, and 3b and c were analysed via a MANOVA analysis. This MANOVA consisted out of the variables autonomous and controlled motivation. This resulted in a not statistically significant difference in locus of motivation based on differing gamification elements, $F\ (6, 202) = 1.51$, $p = .177$; Wilk's $\lambda = .916$. Gamification has not been shown to have a significant effect on either autonomous motivation ($F\ (3,102) = .61$, $p = .608$) or controlled motivation ($F\ (3,102) = 1.90$, $p = .134$).

4.3 Performance

Hypotheses 1d, 2d, and 3d were analysed via a MANOVA analysis. This MANOVA consisted of the variables recognition total test score, average number of mistakes made, uncertainty, and time to completion. This resulted in a statistically insignificant difference in performance based on differing gamification elements, $F\ (12,262.22) = 1.75$, $p = .057$; Wilk's $\lambda = .816$, partial $\eta^2 = .07$. Gamification has not been shown to have a significant effect on the recognition test score ($F\ (3,102) = .53$, $p = .660$), number of mistakes made ($F\ (3,102) = .47$, $p = .703$), or on the time needed to complete the assembly task ($F\ (3,102) = .147$, $p = .931$). However, gamification does have a significant effect on the level of uncertainty of the participants ($F\ (3,102) = 5.50$, $p = .002$, partial $\eta^2 = .14$).

Conducting comparisons via the Bonferroni post hoc correction showed that a leaderboard resulted in significantly lower levels of uncertainty as compared to no gamification elements used (2.0 vs 5.7, $p = .008$), levels (2.0 vs 5.1, $p = .035$), and badges (2.0 vs 6.0, $p = .003$). Therefore, we can conclude that the gamification element leaderboard results in a significantly lower level of uncertainty.

4.4 Mediation Effect

Hypotheses 1e, 2e, and 3e are about exploring the possible mediation effect of competence satisfaction on motivation and performance. As there is a main effect between gamification and the level of uncertainty, a measurement of performance, a mediation analysis has been conducted by making use of the process macro plugin by Hayes (Hayes, 2017). Four steps should be followed to be able to determine if a mediation effect exists or not (Hayes & Preacher, 2014). The first step consists of confirming that the independent variable affects the dependent variable. The goal of the second step is to confirm that the independent variable affects the mediator. The third step is to confirm if the mediator continues to have a significant effect on the dependent variable while accounting for the independent variable. The fourth and last step is about confirming that the independent variable has no significant effect or a significant reduction of the effect on the dependent variable while controlling for the mediator. In the first step of the mediation analysis, the direct effect of gamification on level of uncertainty was significant, b = -1.03, t(104) = -2.85, $p = .005$. However, the effect of gamification on the mediator, competence satisfaction, is not significant, b = -.02, t(104) = -.33, $p = .744$, meaning there is no mediation. The third step of the mediation analysis does, however, show that competence satisfaction, while controlling for gamification, predicts uncertainty level, b = -1.42, t(103) = -2.69, $p = .008$. As there is no mediation effect, it makes no sense to continue with the fourth step.

4.5 Past Experience

The last hypotheses 4a and 4b have been analysed via a MANOVA analysis. This MANOVA exists of the variable past experience with AR and past experience with an assembly task as independent variables, and the performance variables, known as recognition total test score, the average number of mistakes made, uncertainty, and time to completion, served as the dependent variables. This resulted in a not significant effect for the main effect of past experience with AR (F

(16, 263.37) = .86, *p* = .612; Wilk's λ = .855), in a not significant effect for the main effect of past experience with an assembly task (*F*(16, 263.72) = .89, *p* = .579; Wilk's λ = .851), and in a not significant interaction effect between past experience with AR and past experience with an assembly task (*F*(32, 318.75) = .86, *p* = .688; Wilk's λ = .737). As all Wilk's λ statistics are not significant, further description of the MANOVA test statistics does not make sense in this case.

Table 11 gives a brief overview of all the hypotheses and the found results.

Table 11
Overview of the Hypotheses and Their Results

Hypothesis	Result
H1a: The use of badges is positively related to students’ competence satisfaction in completing the task in comparison with the no-manipulations condition.	Rejected
H1b: The use of badges will positively influence the level of autonomous motivation in comparison with the no-manipulations condition.	Rejected
H1c: The use of badges will negatively influence the level of controlled motivation in comparison with the no-manipulation conditions.	Rejected
H1d: The use of badges is positively related to students’ performance in comparison with the no-manipulations condition.	Rejected
H1e: The effect of badges on motivation and performance is mediated by competence satisfaction.	Rejected
H2a: The use of a leaderboard is positively related to students’ competence satisfaction in completing the task in comparison with the no-manipulations condition.	Rejected
H2b: The use of a leaderboard will negatively influence the level of autonomous motivation in comparison with the no-manipulations condition.	Rejected
H2c: The use of a leaderboard will positively influence the level of controlled motivation in comparison with the no-manipulations condition.	Rejected
H2d: The use of a leaderboard is negatively related to students’ performance in comparison with the no-manipulations condition.	Mostly rejected
H2e: The effect of a leaderboard on motivation and performance is mediated by competence satisfaction.	Rejected
H3a: The use of levels is positively related to students’ competence satisfaction in completing the task in comparison with the no-manipulation condition.	Rejected
H3b: The use of levels will negatively influence the level of autonomous motivation in comparison with the no-manipulations condition.	Rejected
H3c: The use of levels will positively influence the level of controlled motivation in comparison with the no-manipulations condition.	Rejected
H3d: The use of levels is negatively related to students’ performance in comparison with the no-manipulations condition	Rejected
H3e: The effect of levels on motivation and performance is mediated by competence satisfaction.	Rejected
H4a: Prior experience with Augmented Reality will positively influence the performance of the participants.	Rejected
H4b: Prior experience with an assembly task will positively influence the performance of the participants.	Rejected

5. Discussion

The present study aimed to determine if and to what extent gamification elements influence motivation and performance when being utilized within AR. Gamification can increase motivation and performance, depending on the situation (Cerasoli et al., 2014; Hamari et al., 2014; Mekler et al., 2017; Seaborn & Fels, 2015), where AR is known for similar effects (Akçayır & Akçayır, 2017; Garzón et al., 2019; Radu, 2014; Sommerauer & Müller, 2018). The present study aimed to combine this strategy and technology to observe if the combination was capable of influencing motivation and performance of university students to an even higher degree. Contrary to what was expected, the results of this study show that the use of gamification elements (levels, badges, and a leaderboard) did not influence individuals' levels of competence satisfaction, locus of motivation, and performance. The variables past experience were expected to influence the performance of the participant, although the statistical analysis showed that this was not the case.

5.1 Competence Satisfaction

For every gamification element, it was expected that competence satisfaction would be influenced. This, however, turned out not to be the case. In general, this lack of an effect is interesting for the following reasons.

First, AR is not a brand-new technology, but in present day, it is still not common. For example, over half of the participants had not used AR before. Nonetheless, this makes the lack of any effect even more interesting because, in general, performance increases whenever a person uses something new and novel. This is known as the novelty effect and has been mentioned as a possible influencer in different papers (Hamari et al., 2014). This novelty effect has not occurred in the present study. This is odd because, as stated earlier, over half of the participants had no prior experience with AR, let alone experience with gamification within AR. It is certainly possible that because the participating university students are associated with a technical university, the students have an affinity with technology, which could have mitigated a possible novelty effect.

Secondly, there was no significant difference in the level of competence satisfaction between all conditions. This can be interpreted as such that competence satisfaction has not been influenced whatsoever. However, when looking at the average level of competence satisfaction on a scale of one to five, we can conclude that it is rather high. By contrast, the level of autonomy satisfaction is similar over all the conditions and has a rather low average. The focus of this study was on manipulating competence satisfaction, because current literature states that AR can make participants more motivated (Akçayır & Akçayır, 2017; Garzón et al., 2019), where the best-educated guess for an explanation is that AR increases autonomy satisfaction (Buchner & Zumbach, 2018; Garzón et al., 2019). However, when looking at the average scores, it becomes apparent that the participants did not feel particularly autonomous in the task they executed. This is an interesting finding, because it seemed to be a key characteristic of AR to increase the autonomy satisfaction of individuals, which seemed rationally plausible. As autonomy and competence satisfaction has only been measured post-test, it is difficult to ascribe the rather high competence satisfaction averages to the use of AR. Nonetheless, it does raise the question if AR influences competence satisfaction rather than autonomy satisfaction.

Thirdly, it is possible that gamification elements simply do not influence competence satisfaction. In theory, they should, but a study done by Mekler et al. (2017) found no significant difference between the used gamification elements and competence satisfaction. Mekler et al. (2017) hypothesized that the lack of this effect was due to the gamification element not offering enough feedback in terms of what kind of performance could be labelled as good or bad. This is something that also exists in the present study, as in the case of the leaderboard the participants never saw their score, were not aware of how many points would be deemed competent, and never were told what time to completion could be deemed as short or long.

5.2 Locus of Motivation

No significant difference between autonomous and controlled motivation has been found. As autonomous motivation is a measure of intrinsic motivation and controlled motivation is a measure of extrinsic motivation, this means that all participants have been equally intrinsically and extrinsically motivated after the experiment. This makes sense, as almost none of the performance indicators have been influenced. It was expected that, whenever an individual was more intrinsically motivated, they would perform better across all performance measurements. The opposite was expected when an individual was more extrinsically motivated. If none of the measurements were influenced, finding no difference between all conditions and their motivational levels would make sense. However, in the case of the leaderboard condition, participants scored significantly lower in terms of uncertainty. Meanwhile, their motivational levels are similar to that of the other conditions. This raises the question what influenced the level of uncertainty of the participants. As a change in gamification element only influenced the level of uncertainty and nothing else, it can be assumed that uncertainty is not influenced by someone’s motivational level.

5.3 Performance

It was hypothesized that gamification would influence the recognition ability of the participants regarding the assembly task, the number of mistakes being made, the level of uncertainty, and the time needed to complete the assembly task. However, only the level of uncertainty is significantly lower in the leaderboard condition as opposed to all other conditions. Uncertainty was the average of the number of times an individual replayed their current step or went back a step. Every replay or step back also resulted in minus points and, thus, would most likely result in a lower place on the leaderboard. It is, therefore, most plausible that the participants in the leaderboard condition wanted to minimize the number of points they could lose by using the replay and step back buttons as little as possible, as they were briefed that this would influence their final score. At the same time, none of the other performance indicators were affected even though they confirmed the steps they needed to take less often than participants in other conditions. This gives the belief that having a clear goal that participants want to achieve, i.e. a spot on the leaderboard, in combination with a mechanism that punishes doubting one’s actions, resulted in either participants wanting to minimize the number of points they could lose while playing, or in individuals that were more sure of their actions, or a combination of both.

The other performance indicators have not been influenced by the leaderboard, while it would have made sense. Needing more time to complete the assembly task also resulted in a penalty that made it more challenging to achieve a spot on the leaderboard. Therefore, it was also expected that individuals in this condition would work faster so as not to lose any points. As every condition had a similar average time to completion, the task itself and the guidance provided via the AR application could be optimized to the point that participants could not work faster without having previous experience with the specific task at hand.

The recognition test investigated the educational ability of the AR application, which was not known to the participant before they started. It is possible that all the individuals performed similarly due to not actively trying to remember the assembly steps, but merely copying them. It has to be noted that all participants only did the assembly task once. Typically, if a person wanted to learn a new task, they would perform the task multiple times. It would be interesting, and a recommendation for future research, to either let participants execute a task multiple times before testing their learning ability or to test the participants' knowledge every time they completed the assembly task to see if progress can be measured. It is expected that the mistakes made will also decrease after every successive time the participant performs the task.

The other gamification elements did not influence the performance indicators at all. It is possible that the goal for the levels and badges condition was not clear or exact enough, as Jung et al. (2010) found that individuals perform better when they have a specific goal, as opposed to being asked to do their best. This raises the question if, in the badges condition, being asked if the participants can find them all is actually asking them to do their best or giving them a specific goal. In this instance, it is most likely that the participants interpreted the question as non-committal. It might have helped if it was clearly stated that it was the goal of the participant to find all the badges. That way, they would have been sent on a mission, instead of having a question, which they could simply ignore and continue with the assembly task. The participant did receive all the badges regardless of whether they were looking for them or not. The badges were programmed to show up at specific moments that the participant was not aware of.

In terms of the levels condition, the participants were explicitly told that it was their goal to complete all the levels. They were also made aware of when they finished a level and started the next one. It is possible, however, that the levels either had too similar of an objective so that the challenge diminished or that the levels did not have the correct build-up in difficulty. Every level consisted of putting parts together, which results in the same objective no matter what level the person was in. Mekler et al. (2017) used a progress bar that indicated how far along the person was in achieving the next level and how much points they still needed to gather. It is plausible that the participants in the present study had no clear sense of progression, which would have been avoided if a progress bar had been used.

Another possibility is that the levels did not complement each other in terms of difficulty. For example, level two only consisted of three steps. This decision was made because, generally, people have difficulty completing the steps due to the screws barely fitting and using counterclockwise rifling. However, due to only having a few steps, it is possible that participants felt like it was not on par with the other levels that all consisted out of at least twice as many steps. This could have created a disbalance between the levels because the other levels felt finished after doing the portrayed steps, whereas the second level could have surprised the individual by only taking three steps.

5.4 Mediation Effect

As gamification has a main effect on the level of uncertainty of the participant, it was relevant to see if this main effect was mediated by the level of competence satisfaction. It was found early on that competence satisfaction was not predicted by the independent variable gamification and, thus, a mediation effect could not exist. Nonetheless, it was observed that competence satisfaction had some sort of influence on the level of uncertainty. This means that at first glance, more competence satisfaction would result in less uncertainty. However, the directionality of this observation is most likely the other way around. As gamification, the only construct manipulated in the present study did not have a significant effect on the level of competence satisfaction, it is more reasonable to think that gamification influenced level of uncertainty, which, in turn, influenced the level of competence satisfaction. When we discussed performance, it was stated that participants most likely wanted to minimize their loss of points or were more certain of their actions due to the use of a leaderboard. When we combine this with the aforementioned effect of uncertainty on competence satisfaction, it is either the case that wanting to minimize the points you can lose makes an individual, unbeknownst, feel more competent, or that being more sure of one’s actions influences one’s level of competence satisfaction. As the present study cannot answer which of the two options is the case, this immediately makes it a recommendation to research this further to uncover the complete effect a leaderboard can have.

5.5 Past Experience

The variables past experience with AR, level of liking of those past experiences, and past experience with an assembly task did have no significant effect on the performance of the participants. It was reasonable to hypothesize that this effect would occur, as transferring prior experience to a new situation has been known to have an effect (Nokes, 2009). This avenue was taken into account because Hamari et al. (2014) concluded in their literature review that not taking past experiences into account when dealing with AR is a methodological limitation of the papers they observed. More recent literature reviews also shows that past experience was still not a measure that was taken into account when researching the effects AR can have (Akçayır & Akçayır, 2017; Garzón et al., 2019; Garzón & Acevedo, 2019). The present study is a start for understanding the possible effect past experiences with AR can have when people make use of AR and gamification. It is a possibility that the self-created questionnaire to gauge the level of past experience of the participants is a methodological shortcoming of the present study. Hamari et al. (2014) also mentioned that when previous studies measured past experience, it was often by making use of non-standardized and non-validated scales.

5.6 Practical Implications

None of the gamification elements had an observed effect on competence satisfaction. This indicates that educational systems and businesses should wait until future research has been conducted that has observed any effects. Until then, using gamification elements to boost the competence satisfaction of users will only take more time and costs more, without any observed benefits.

As motivation can play an essential factor in whether employees and students will act in certain ways or do certain things, being able to manipulate it is very relevant. As the present study has not found gamification being able to help with making participants more intrinsically motivated, it is implied that adding those elements to an educational AR solution would not hold any merit. Until it has been observed that gamification within AR can increase motivation, it does not make sense to apply gamification elements to real-world solutions.

In terms of performance, it is found that a leaderboard can significantly reduce the level of uncertainty, while not influencing any other performance indicator in a significant way. As this effect is found with students, the question arises if this effect can be utilized with other groups as well. When having a hands-on situation that is new for most students, using a leaderboard can help the students in trusting their actions without going back and double-checking. For businesses, the same mechanism can work as well, to stimulate employees in trusting what they saw and what they think they should do, as opposed to double-checking the instructions. However, the effect has only been found concerning students, so businesses are better off waiting until the same effect has been observed in a business-like setting.

5.7 Limitations and Future Research

In terms of competence satisfaction, it is, firstly, possible that the way AR has been used influenced the efficacy of this technology. AR is known for the freedom it gives due to it acting within 3D space, while in this study, participants were seated behind a desk. It was a designated lab area, which means that it was not possible to change the set-up easily. The participants were free to stand up and use the mobile AR application any way they deemed fit. However, most of the participants chose to stay mostly seated. Assuming that AR can indeed enhance motivation, the stationary way of working could be the reason why no hypothesized effect was found. This does not mean that the assembly task used in this study is not sufficiently compatible with AR. It means that a lab setting with, for example, a stand-up desk or the ability to walk around the 3D space could have made the use of AR more effective. Therefore, it is possible that if the environment stimulated the participant to fully use their autonomy, that AR would have been more effective in enhancing autonomy satisfaction. This is something to take note when performing AR-related research in the future.

Secondly, it is possible that AR can improve motivation, however, as discussed earlier, the ability to enhance autonomy satisfaction could be wrongly attributed. It has been stated in earlier research and meta-analyses that AR most likely influenced motivation due to it enhancing the participants' autonomy. This seems logical and would fit the technology, but it has not yet been confirmed. AR's ability to enhance motivation may have been observed due to a novelty effect associating the technology. This can be seen as another limitation because the body of research on AR and the inner workings of motivation is rather young.

Thirdly, the way the gamification elements have been implemented could have negated the effect they can have. Badges, for example, were created based on the definition of Hakulinen et al. (2015). However, Montola et al. (2009) and Hamari and Eranti (2011) both define badges as an addition to a process that results in non-compulsory objectives. The present study presented the badges with the question if the participant could find them all. It was thought that this set-up gave the participant enough of a stimulating goal and would make them a non-compulsory objective. However, the participants only needed to progress through the task linearly to find the badges. There were no subtasks they needed to complete to achieve a badge. The participants did not need to work for or be motivated to find the badges. This could explain the lack of an effect. This results in the recommendation for future research to test the effect of badges when they can only be achieved when the user actively pursues them. This could help define a unified, evidence-based definition of what makes badges effective.

In the case of levels, it is possible that their implementation was too simple. After every central part was assembled, a level was finished. Generally, when a level finishes, a player goes to the next stage. In many games, the next stage takes place in the same setting, but the requirements are different. In the present study, it could be that the requirements of every level were too similar for the participant to feel like they were working through different levels. Consequently, it is advisable to explore in future research if a task with levels that differ in requirements affects competence satisfaction. This could not be achieved in the current study due to being constrained by what was possible with the piece of equipment used in the assembly task.

When looking at the locus of motivation, it is unknown if the participants' motivation was increased, decreased, or unchanged as no pre-test was conducted. This was not done because it was expected that the gamification elements would have some effect on motivation. Therefore, a notable difference was expected between the none and the gamification conditions. As this did not turn out to be the case, not having done a pre-test appears to be a limitation of this study. Consequently, it is recommended to perform a pre-test to be able to make sure that the created manipulations at least influence the variables as expected.

Furthermore, the level of uncertainty was influenced by the leaderboard condition. As it was expected that this would only happen when the participant would feel more intrinsically motivated as opposed to extrinsically motivated, it is a recommendation that future research validates that uncertainty, as used in the present study, is not influenced by one's motivational level. It might be that feeling less uncertain makes individuals feel more competent, whereas feeling more competent and more autonomous does not influence the level of uncertainty an individual experiences.

All two variables concerning past experience had no significant influence on performance. The use of a self-made scale to measure those variables may be a limitation to be able to effectively measure the effect past experience can have. It is, therefore, a recommendation for future research to either use an already validated scale or validate the scales that are used to make sure that any measured effect of past experience is done in a standardized way.

5.8 Conclusion

This research explored the interaction between AR, gamification, and motivation as one of the first of its kind. A leaderboard, badges, and levels were added to a mobile AR application that instructed the participant how to assemble a piece of equipment. The goal was to determine if adding gamification elements could improve motivation and, ultimately, the performance of the participants. In addition to this, it was explored if past experience with AR and an assembly task would influence the performance of the participants. It was concluded that none of the gamification elements influenced motivation, and none of the past experience variables influenced motivation and performance. The leaderboard condition, however, did influence the level of uncertainty of the participants, while not influencing any of the other indicators for performance. This is not in line with what was expected and previously found. This brings us to the conclusion that more empirical research is needed to determine when and how gamification influences motivation and performance, if the combination of AR and gamification is beneficial or not, and through what way manual instructions can be best digitalized. Even though the present study has not found evidence for almost all of their stated hypotheses, it does raise interesting new questions, and it is believed that the study functions as a first step in the direction to combine gamification, AR, and motivation in an empirical way.

6. References

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Altinpulluk, H. (2019). Determining the trends of using augmented reality in education between 2006-2016. *Education and Information Technologies*, *24*(2), 1089–1114. https://doi.org/10.1007/s10639-018-9806-3

Aparicio, A. F., Sánchez, J. L. G., Gutiérrez-Vela, L. F., & Isla-Montes, L. F. (2012). Analysis and application of gamification. *Proceedings of the 13th International Conference on Interacción Persona-Ordenador*, *Article 17*, 1–2. https://doi.org/10.1145/2379636.2379653

Badri, R., Amani-Saribaglou, J., Ahrari, G., Jahadi, N., & Mahmoudi, H. (2014). School culture, basic psychological needs, intrinsic motivation and academic achievement: Testing a casual model. *Mathematics Education Trends and Research*, *2014*(3), 1–13. https://doi.org/10.5899/2014/metr-00050

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, *71*, 550–585. https://doi.org/10.1016/j.chb.2016.08.049

Bottani, E., & Vignali, G. (2019). Augmented reality technology in the manufacturing industry: A review of the last decade. *IISE Transactions*, *51*(3), 1–64. https://doi.org/10.1080/24725854.2018.1493244

Buchner, J., & Zumbach, J. (2018). *Promoting intrinsic motivation with a mobile augmented reality learning environment* (ED590357). ERIC. https://eric.ed.gov/?id=ED590357

Buckley, P., & Doyle, E. (2016). Gamification and student motivation. *Interactive Learning Environments*, *24*(6), 1162–1175. https://doi.org/10.1080/10494820.2014.964263

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*(4), 980–1008. https://doi.org/10.1037/a0035661

Chen, B., Van Assche, J., Vansteenkiste, M., Soenens, B., & Beyers, W. (2015). Does psychological need satisfaction matter when environmental or financial safety are at risk? *Journal of Happiness Studies*, *16*(3), 745–766. https://doi.org/10.1007/s10902-014-9532-5

Chen, Y.-C. (2006). A Study of Comparing the Use of Augmented Reality and Physical Models in Chemistry Education. *Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications*, 369–372. https://doi.org/10.1145/1128923.1128990

Dahlstrøm, C. (2017). *Impacts of gamification on intrinsic motivation*. Norway: Norwegian University of Science and Technology, Department of Design. Retrieved from university 'Norwegian University of Science and Technology' website: https://www.ntnu.edu/design/student-articles.

Deci, E. L. (1972). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology*, *22*(1), 113–120. https://doi.org/10.1037/h0032355

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press.

Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deci, E. L., & Ryan, R. M. (2002). *Handbook of Self-Determination Research* (1st ed.). University of Rochester Press.

Denny, P. (2013). The effect of virtual achievements on student engagement. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 763. https://doi.org/10.1145/2470654.2470763

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining “Gamification.” *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Dooley, David., & Vos, H. J. (2008). *Social research methods* (4th edition.). Pearson Custom Publ; WorldCat.

Dresel, M., & Grassinger, R. (2013). Changes in Achievement Motivation among University Freshmen. *Journal of Education and Training Studies*, *1*(2), 159–173. https://doi.org/10.11114/jets.v1i2.147

Erbas, C., & Demirer, V. (2019). The effects of augmented reality on students’ academic achievement and motivation in a biology course. *Journal of Computer Assisted Learning*, *35*(3), 450–458. https://doi.org/10.1111/jcal.12350

Gagné, M., Forest, J., Gilbert, M.-H., Aubé, C., Morin, E., & Malorni, A. (2010). The Motivation at Work Scale: Validation Evidence in Two Languages. *Educational and Psychological Measurement*, *70*(4), 628–646. https://doi.org/10.1177/0013164409355698

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of Augmented Reality on students’ learning gains. *Educational Research Review*, *27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.001

Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality 23*, 447–459. https://doi.org/10.1007/s10055-019-00379-9

Gentner, D. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 99–129). Erlbaum.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Gilliland, S. W., & Landis, R. S. (1992). Quality and Quantity Goals in a Complex Decision Task: Strategies and Outcomes. *Journal of Applied Psychology*, *77*(5), 672–681. https://doi.org/10.1037/0021-9010.77.5.672

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 233–240. https://doi.org/10.1037/a0012758

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*(3), 175–213. https://doi.org/10.1023/A:1005614228250

Hakulinen, L., Auvinen, T., & Korhonen, A. (2015). The Effect of Achievement Badges on Students' Behavior: An Empirical Study in a University-Level Computer Science Course. *International Journal of Emerging Technologies in Learning (IJET)*, *10*(1), 18–29. https://doi.org/10.3991/ijet.v10i1.4221

Hamari, J., & Eranti, V. (2011). Framework for Designing and Evaluating Game Achievements. *Proceedings of DiGRA 2011 Conference: Think Design Play*, *6*.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. *Proceedings of the 2014 47th Hawaii International Conference on System Sciences*, 3025–3034. https://doi.org/10.1109/HICSS.2014.377

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Publications.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470. https://doi.org/10.1111/bmsp.12028

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. J. Holyoak & J. A. Barnnden (Eds.), *Analogical connections.* (pp. 113–180). Ablex Publishing.

Hornstra, L., Weijers, D., van der Veen, I., & Peetsma, T. (2016). *Het motiveren van leerlingen met verschillende prestatieniveaus en achtergrondkenmerken* [Motivating students with different performance levels and background characteristics]. (Report No. 405-15–514). University of Utrecht.

Hung, Y.-H., Chen, C.-H., & Huang, S.-W. (2017). Applying augmented reality to enhance learning: A study of different teaching materials: Augmented reality enhances learning. *Journal of Computer Assisted Learning*, *33*(3), 252–266. https://doi.org/10.1111/jcal.12173

Inspection of Education. (2020). *Staat van het Onderwijs 2020* [State of Education 2020]. Ministry of Education, Culture, and Science. https://www.onderwijsinspectie.nl/documenten/rapporten/2020/04/22/staat-van-het-onderwijs-2020

Jung, J. H., Schneider, C., & Valacich, J. (2010). Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. *Management Science*, *56*(4), 724–742. https://doi.org/10.1287/mnsc.1090.1129

Khan, W. A., Raouf, A., & Cheng, K. (2011). *Virtual Manufacturing*. Springer London. https://doi.org/10.1007/978-0-85729-186-8

Krevelen, D. W. F. (Rick) van, & Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality*, *9*(2), 1–20. https://doi.org/10.20870/IJVR.2010.9.2.2767

Kühne, C. (2019). *Balance in the satisfaction of basic psychological needs (SDT) across life domains and its relation to subjective wellbeing in university students* [Bachelor Thesis, University of Twente]. http://purl.utwente.nl/essays/78258

Kumar, J. (2013). Gamification at work: Designing engaging business software. In A. Marcus (Ed.), *Design, user experience, and usability. Health, learning, playing, cultural, and cross-cultural user experience* (pp. 528–537). Springer Berlin Heidelberg.

Lepper, M. R., & Greene, D. (1978). *The hidden costs of reward: New perspectives on the psychology of human motivation.* Lawrence Erlbaum.

Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of education and learning: A review of empirical literature. *Proceedings of the 2nd International GamiFIN Conference*, *2186*, 11–19. CEUR-WS.

McDaniel, R., Lindgren, R., & Friskics, J. (2012). Using badges for shaping interactions in online learning environments. *Proceedings of the 2012 IEEE International Professional Communication Conference*, 1–4. https://doi.org/10.1109/IPCC.2012.6408619

Meijers, F. (2006). Loopbaanbegeleiding in de beroepskolom: Tussen droom en daad [Career guidance in the professional segment: Between dreaming and doing]. *Pedagogiek*, *26*(1), 26–44.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, *71*, 525–534. https://doi.org/10.1016/j.chb.2015.08.048

Montola, M., Nummenmaa, T., Lucero, A., Boberg, M., & Korhonen, H. (2009). Applying game achievement systems to enhance user experience in a photo sharing service. *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, 94–97. https://doi.org/10.1145/1621841.1621859

Morris, B. J., Croker, S., Zimmerman, C., Gill, D., & Romig, C. (2013). Gaming science: The “Gamification” of scientific thinking. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00607

Nokes, T. J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning*, *15*(1), 1–36. https://doi.org/10.1080/13546780802490186

Núñez, J. L., & León, J. (2015). Autonomy support in the classroom: A review from self-determination theory. *European Psychologist*, *20*(4), 275–283. https://doi.org/10.1027/1016-9040/a000234

Pellas, N., Fotaris, P., Kazanidis, I., & Wells, D. (2018). Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Reality*, *23*(4), 329–346. https://doi.org/10.1007/s10055-018-0347-2

Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for Motivation, Engagement and Wellbeing in Digital Experience. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00797

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*(4), 734–746. https://doi.org/10.1037/0022-0663.99.4.734

Reeves, B., & Read, J. L. (2009). *Total engagement: Using games and virtual worlds to change the way people work and businesses compete*. Harvard Business School Press.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066x.55.1.68

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, *30*(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Saltan, F., & Arslan, Ö. (2017). The use of augmented reality in formal education: A scoping review. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(2), 503–520. https://doi.org/10.12973/eurasia.2017.00628a

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Soenens, B., Sierens, E., Vansteenkiste, M., Dochy, F., & Goossens, L. (2012). Psychologically controlling teaching: Examining outcomes, antecedents, and mediators. *Journal of Educational Psychology*, *104*(1), 108–120. https://doi.org/10.1037/a0025742

Sommerauer, P., & Müller, O. (2018). Augmented reality for teaching and learning—A literature review on theoretical and empirical foundations. *Proceedings of the 26th European Conference on Information Systems*, 1–17. https://aisel.aisnet.org/ecis2018_rp/31/

Sun, H., Li, W., & Shen, B. (2017). Learning in physical education: A self-determination theory perspective. *Journal of Teaching in Physical Education*, *36*(3), 277–291. https://doi.org/10.1123/jtpe.2017-0067

Tröndle, M. (2016). *Gamification – Does it lead to higher motivation?* [Bachelor Thesis, Hochschule Furtwangen]. https://nbn-resolving.org/urn:nbn:de:bsz:fn1-opus4-19618

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. In T. C. Urdan & S. A. Karabenick (Eds.), *The decade ahead: Theoretical perspectives on motivation and achievement (Advances in motivation and achievement)* (Vol. 16 Part A, pp. 105–165). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-7423(2010)000016A007

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, *23*(3), 263–280. https://doi.org/10.1037/a0032359

Vansteenkiste, M., Sierens, E., Soenens, B., & Lens, W. (2007). Willen, moeten en structuur: Over het bevorderen van een optimaal leerproces[To want, to have and structure: On facilitating an optimal learning process]. *Begeleid Zelfstandig Leren*, *16*(2), 37–58.

Wang, X., Schneider, C., & Valacich, J. S. (2015). Enhancing creativity in group collaboration: How performance targets and feedback shape perceptions and idea generation performance. *Computers in Human Behavior*, *42*, 187–195. https://doi.org/10.1016/j.chb.2014.02.017

Witte, K. D., Cabus, S., & Groot, W. (2014). *Voortijdig schoolverlaten: Rapportage van TIER/Platform31* [Early school leaving: Report from TIER/Platform31]. Platform31. https://www.platform31.nl/uploads/media_item/media_item/23/81/voortijdig_schoolverlaten-1400572255.pdf

Yoon, S. A., Elinich, K., Wang, J., Steinmeier, C., & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*, *7*(4), 519–541. https://doi.org/10.1007/s11412-012-9156-x

7. Appendices

Appendix 1a – SIMS scale

Why were you engaged in the task just performed?

	Strongly disagree	disagree	More or less disaggree	Undecided	More or less agree	Agree	Strongly agree
1. Because I think that this activity is interesting	○	○	○	○	○	○	○
2. Because I am doing it for my own good	○	○	○	○	○	○	○
3. Because I am supposed to do it	○	○	○	○	○	○	○
4. There may be good reasons to do this activity, but personally I don't see any	○	○	○	○	○	○	○
5. Because I think that this activity is pleasant	○	○	○	○	○	○	○
6. Because I think that this activity is good for me	○	○	○	○	○	○	○
7. Because it is something that I have to do	○	○	○	○	○	○	○
8. I do this acitvity but I am not sure if it is worth it	○	○	○	○	○	○	○
9. Because this activity is fun	○	○	○	○	○	○	○
10. By personal decision	○	○	○	○	○	○	○
11. Because I don't have any choice	○	○	○	○	○	○	○
12. I don't know; I don't see what this activity brings me	○	○	○	○	○	○	○

13. Because I feel good when doing this activity	○	○	○	○	○	○	○
14. Because I believe that this activity is important for me	○	○	○	○	○	○	○
15. Because I feel that I have to do it	○	○	○	○	○	○	○
16. I do this activity, but I am not sure it is a good thing to pursue it	○	○	○	○	○	○	○

Appendix 1b – BPNSAF scale

1	2	3	4	5
Not true at all				Completely true

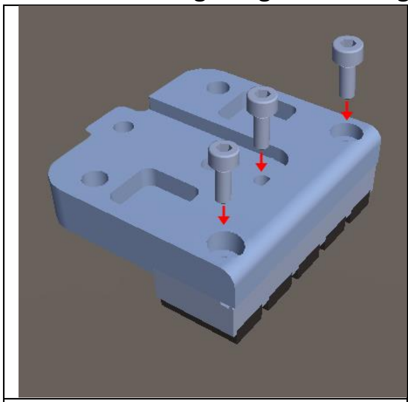
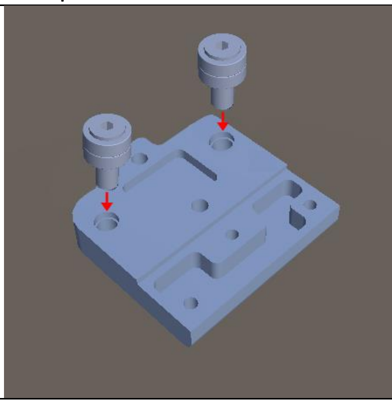
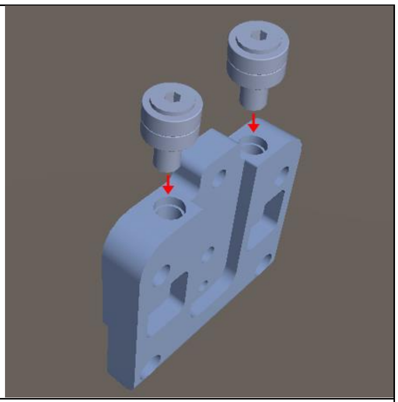
During the just performed task…						
1.	…I felt a sense of choice and freedom.	1	2	3	4	5
2.	…I felt confident that I could do well.	1	2	3	4	5
3.	…I felt that my decisions reflected what I really wanted.	1	2	3	4	5
4.	…I felt capable at what I was doing.	1	2	3	4	5
5.	…I felt my choices expressed who I really am.	1	2	3	4	5
6.	…I felt competent to achieve my goals.	1	2	3	4	5
7.	…I felt I was doing what really interests me.	1	2	3	4	5
8.	…I felt I can successfully complete difficult tasks.	1	2	3	4	5

Autonomy need satisfaction: 1,3,5,7
Competence need satisfaction: 2,4,6,8

Appendix 1c – Recognition test

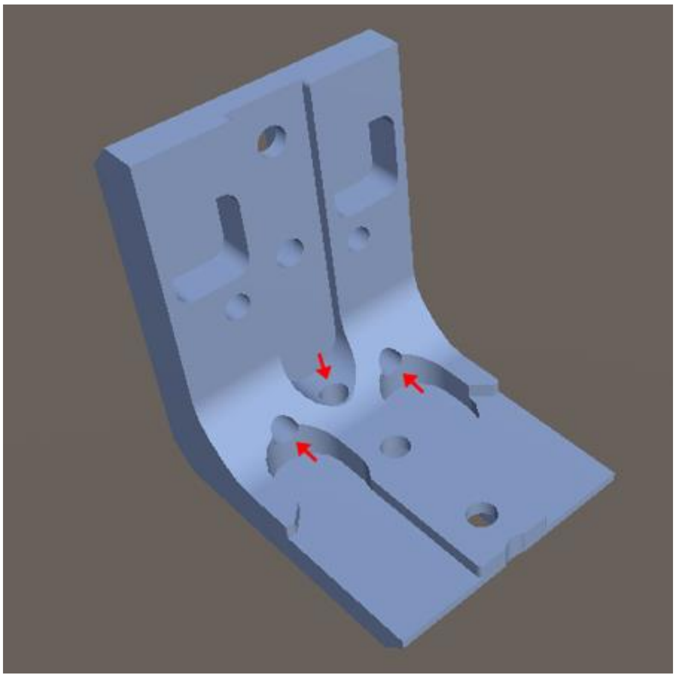
Question 1 – sequence

Put the following images in the right sequence:

		
a	b	c

Question 2 – tool

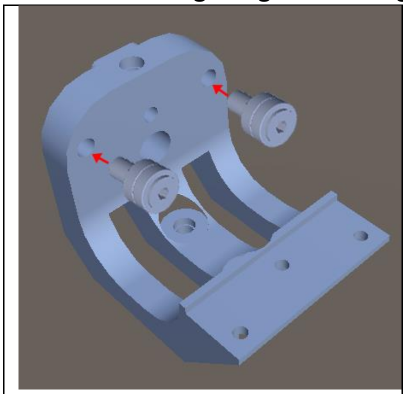
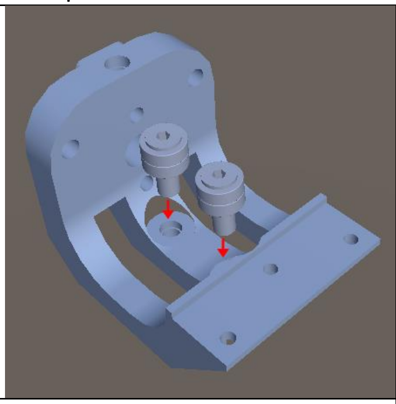
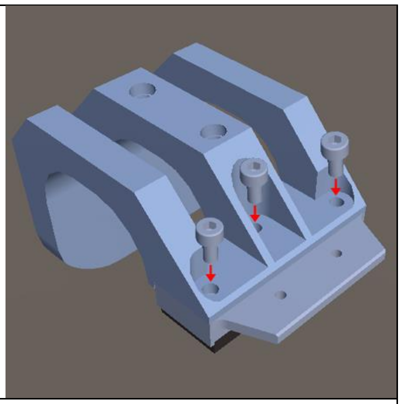
Which tool should you use to screw the pointed at bolts down?



a) Allen key 3
b) Torx key 25
c) Allen key 2.5
d) Torx key 35

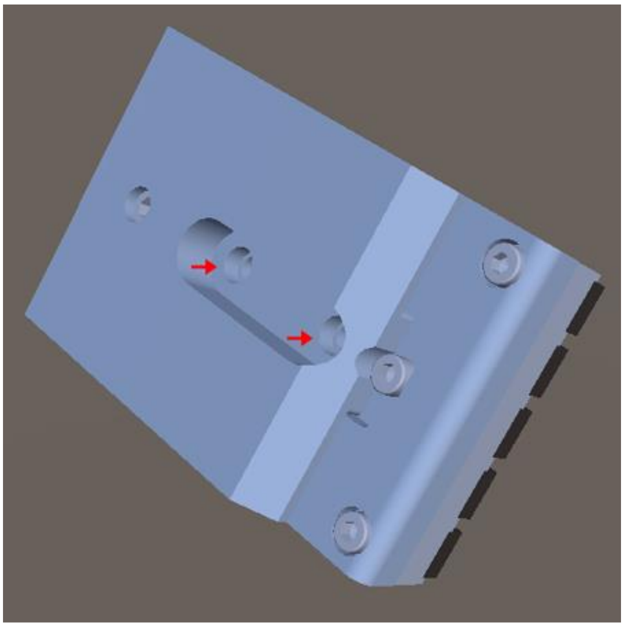
Question 3 – sequence

Put the following images in the right sequence:

		
b	a	c

Question 4 – tool

Which tool should you use to screw the pointed at bolts down?



a) Torx key 25
b) Torx key 35
c) Allen key 2.5
d) Allen key 3

Appendix 1d – Demographic information

What is your age?		
What is your gender?		
Male	Female	Other
What study do you follow?		

How experienced would you rate yourself in the use of AR applications?				
1	2	3	4	5
Not at all				Completely true

If you had any previous experiences with AR, how well did you like those experiences?				
1	2	3	4	5
Not at all				Completely true

How experienced would you rate yourself in the type of task just performed?				
1	2	3	4	5
Never assembled anything				I perform similar tasks every week

Appendix 2 – The consent form

UNIVERSITY OF TWENTE.

Dear sir / madam,

Thank you for participating in this study. This research in being conducted in the context of the Master Communication Studies at the University of Twente. Before you start, I would like to ask you to read the following carefully.

Goal
This research aims to take a closer look at the mechanisms of motivation. The study will take you approximately 45 minutes.

Anonymousity
Participating in this research will be anonymous. The recording and your answers are only available to the research and the supervisor. The collected date will be kept confidential and will not be shared with third parties.

With kind regards,

Egbert Roelofsen (e.g.roelofsen@student.utwente.nl)

I've been sufficiently informed about this investigation. I know that my participation will remain completely anonymous, and that I can stop my participation at any time.

- [ ] yes, I consent to participating in this research.

Appendix 3a – First Principal Component Analysis of the items of the SIMS

Items	Component 1	2	3	4	5	Constructs
5. Because I think that this activity is pleasant.	**0.904**	-0.180	-0.145	0.035	-0.028	Intrinsic Motivation
9. Because this activity is fun.	**0.876**	-0.178	-0.201	0.077	-0.125	
13. Because I feel good when doing this activity.	**0.812**	-0.136	-0.184	0.090	-0.102	
1. Because I think that this activity is interesting.	**0.804**	-0.193	-0.287	0.128	-0.002	
3. Because I am supposed to do it.	-0.190	**0.855**	0.084	-0.033	0.093	Internal Regulation
7. Because it is something that I have to do.	-0.194	**0.847**	0.125	0.020	0.172	
15. Because I feel that I have to do it.	-0.116	**0.794**	0.001	-0.067	0.245	
16. I do this activity, but I am not sure it is a good thing to pursue it.	-0.179	0.133	**0.779**	0.049	0.033	External Regulation
4. There may be good reasons to do this activity, but personally I don't see any.	0.013	0.190	**0.676**	-0.395	0.162	
8. I do this activity but I am not sure if it is worth it.	-0.392	0.027	**0.653**	-0.022	-0.087	
12. I don't know; I don't see what this activity brings me.	-0.383	-0.102	**0.612**	-0.131	0.140	
2. Because I am doing it for my own good.	-0.092	-0.220	0.055	**0.826**	0.074	Amotivation
14. Because I believe that this activity is important for me.	0.245	0.120	-0.280	**0.722**	0.038	
6. Because I think that this activity is good for me.	0.476	0.101	-0.101	**0.636**	-0.241	
11. Because I don't have any choice.	-0.144	0.188	0.063	0.152	**0.881**	Wrong loadings
10. By personal decision.	0.036	-0.333	-0.076	0.184	**-0.788**	

Notes. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Loadings larger than 0.50 are highlighted in bold.

Appendix 3b – Principal Component Analysis of the Aut and Comp items of the BPNS scale

Items	Component		Constructs
	1	2	
I felt capable at what I was doing.	**0.874**	-0.044	Competence Satisfaction
I felt confident that I could do well.	**0.828**	0.076	
I felt competent to achieve my goals.	**0.762**	0.178	
I felt I can successfully complete difficult tasks.	**0.612**	0.345	
I felt my choices expressed who I really am.	0.008	**0.798**	Autonomy Satisfaction
I felt that my decisions reflected what I really wanted.	0.136	**0.751**	
I felt a sense of choice and freedom.	0.072	**0.749**	
I felt I was doing what really interests me.	0.387	**0.576**	

Notes. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Loadings larger than 0.50 are highlighted in bold.