Using Game Elements in Audience Response Systems: The Effect of Self-Efficacy on Performance and Gaming Strategies

Jasmine C. Jethu 2739488

Faculty of Behavioural, Management and Social Sciences, University of Twente

M12: Bachelor Thesis 202000384

Supervisor: Dr. P. M. Papadopoulos

2nd Supervisor: L. Hogenkamp

Date: 1-07-2024

Abstract

Background: The gamification element of point wagering regards the user as autonomous and provides a sense of control. The level of self-efficacy of a student could affect the implementation of this gamification strategy in a gamified test (or: quiz) activity.

Objective: This study aims to identify the effect of self-efficacy on the performance and utilization of this gaming strategy in audience response systems (i.e. quiz activities), as well as measure the effects of point wagering on quiz scores.

Methods: 48 students were divided into two groups (control group $n = 26$, treatment group $n = 22$). All participants answered a general knowledge test in Qualtrics, in which the treatment group was exposed to the gamification element of point wagering. Here, students were allowed to distribute a total of four points over four multiple choice answers in a question, in which the most points given to an answer equalled the most shown confidence (variations included *4-0-0-0*, or *3-1-0-0*, or *2-2-0-0*, or *2-1-1-0*, or *1-1-1-1* in any order). Additionally, all students were asked to fill in an academic self-efficacy instrument beforehand.

Results: Analysis showed no significant difference in student's quiz scores among the two groups. Moreover, in the control group, no correlation was found between self-efficacy and the quiz scores. However, in the treatment group, a positive moderate correlation was found between self-efficacy and the quiz scores, as well as between the quiz scores and the confidence shown in how the gamification element was utilized.

Conclusion: Self-efficacy does not affect the quiz scores of students in a normal quiz. However, with regards to the gamification element of point wagering, it appears that self-efficacy and the confidence shown by students are correlated to their quiz scores. Therefore, self-efficacy is an important factor in a quiz with wagering mechanisms.

Keywords: gamification, audience response systems, self-efficacy, point wagering

1. Introduction

In an educational field that is constantly evolving, traditional approaches to teaching have shown their limitations in engaging students effectively, while newer interactive approaches, such as Audience Response Systems (ARS(s)), have shown many benefits like higher retention and enhanced engagement in students, and better performance (Mazur, 2009; Ruhl et al., 1987; Hunsu et al., 2016). Despite its many established benefits, challenges (e.g. student guessing) persist (Nicholson & Bassignani, 2009; Schmidt, 2011). This research suggests that a good solution to address the issues of ARSs is implementing gamification elements, specifically point wagering.

Gamification (or gamified learning) includes the usages of game elements in environments that are non-game like (Deterding et al, 2011). Gamification is effective as the game-like elements hold a high motivational potential, gamified learning is often experienced as more enjoyable (Sailer et al., 2017; Turan et al., 2018). More specifically, the gamification element of points often leads to better performance (Hellberg & Moll, 2018). This research introduces a variation on the typical game element of points, namely point wagering. Like ARSs and gamified learning, academic self-efficacy (i.e. a student's belief in their ability to succeed at a task) is also known to positively correlate with academic performance (Honicke & Broadbent, 2016; Ahmadi, 2020). However, its impact on gamification, particularly point wagering, is underexplored, as a large amount of literature focuses on the effect of gamification on self-efficacy, instead of the effect of self-efficacy on gamification.

This research focuses on filling the research gap by investigating how self-efficacy influences student performance and wagering behaviour in ARS-based quizzes, as well as establishing the effect of point wagering on student performance.

1.1 Traditional versus Interactive Learning Approaches

Mazur (2009) noticed the "traditional approach" of preparing a lecture and teaching from his notes was ineffective. He calls it a "transfer of information" (p. 50.), which included little to no interaction with his students. Over the years, this approach has quickly become outdated due to rapid changes in the educational field. Accordingly, Ruhl et al. (1987) have already compared this traditional approach to newer, more interactive learning approaches. Their results show higher short- and long-term retention in students when using interactive learning methods. Specifically, adding short breaks during interactive lectures to compare study notes helped increase retention (Hunsu et al., 2016). Ultimately, these interactive learning

approaches are characterised by their ability to involve both the teacher and the learner, to encourage learners to participate by giving them a hands-on experience, and to include students in their own learning process (Senthamarai, 2018). Additionally, Mazur concluded that learning almost triples through interactive learning, compared to the traditional approach (Mazur, 2009).

To steer away from the traditional approach identified by Mazur (2009), teachers must take on a different, more interactive approach, in which the control of the learning process is not constantly on the teacher, but also on the student (Gleason et al., 2011). One suggestion for the improvement of quality of lectures and teachings, is the integration of new educational technologies, which involves the 'implementation of relevant tools and processes that enhance teaching processes and facilitate improved learning' (Ng, 2015. p.96.). Ng (2015) believes that an interdependency exists between instruction and technology, and therefore being technically competent in using tools enables educators to be more innovative in their teaching. While considering how students experience these technologies, these technologies must help them reflect, encourage, and engage them in their learning process, which helps to understand the content and identify gaps in their knowledge (Senthamarai, 2018). A technology that fits this description and engages students in their learning process is an Audience Response System (ARS(s)), which allows the teacher to propose questions to a large group of students at once and collect all individual answers (or group responses), through handheld devices (Wood & Shirazi, 2020). For this study, the emphasis is put on ARSs, as they succeed in enhancing the involvement of students and creating active engagement (Hunsu et al., 2016).

1.2 Benefits and Limitations of Audience Response Systems

The so-called Audience Response Systems or voting systems or student response systems (Ruhl et al., 1987) allow all students to be addressed at once (Papadopoulos et al., 2021a). ARS systems provide multiple opportunities to create an interactive learning environment, with different sorts of closed questions, group exercises, and discussions (which can be created and organised by the teacher) to effectively engage the students, while simultaneously providing the teacher with distributions of the scores for feedback purposes (Papadopoulos et al., 2021b). As mentioned before, ARSs succeed in engaging the student, helping them reflect and encourage them. Students claim that ARSs are able to push and motivate the student to think about important concepts, provide them with opportunity to discuss questions with peers, and make learning more fun by using devices (Kay & LeSage, 2009).

It is evident that ARSs are effective, as students who use these systems score significantly higher course grades than their peers (Edmonds & Edmons, 2008; Pradhan et al., 2005). Moreover, for students who only like to share their answers anonymously, ARSs are seen as an approach that diminishes the effects of unsure students who follow the majority, and only high-achieving students raising their hands (Mader & Bry, 2019). Apparently, ARSs show great benefits for the learning processes of students, as they engage the students, allow for discussions, lead to higher grade scores, and push them to think more critically about important topics, while simultaneously providing a sense of protection to students that only like to participate in anonymous classroom activities.

However, students have also reported negative experiences with ARSs. First, students seem to worry about the pace of the lectures, as they fear that including ARSs too frequently slows down its pace, resulting in not covering all the materials needed (Nicholson & Bassignani, 2009). Thus, students believe that ARSs unnecessarily slow down the pace, and simultaneously allow teachers to advance too easily during lectures (Silliman & McWilliams, 2004). Another limitation exists as students admit to guessing while using ARS. For example, some reasons mentioned by students were poor quality of content used in the ARS, or solely guessing because they were unsure (Donohue, 2014; Schmidt, 2011). Schmidt (2011) points out that while using ARSs, teachers must be aware of the latter, as a correct answer does not directly mean an understanding of why it is true. In conclusion, Thomas et al (2011) assert that not all questions are answered by students constantly. They provide several reasons why students might ignore a question: frustration over difficult questions, disengagement due to simple questions, or question fatigue.

1.3 The Use of Gamification Elements

1.3.1 Gamification Elements

One suggestion for tackling guessing in ARS activities could be the use of gamification methods. Games hold a high level of motivational potential and are popular among all age groups. This high motivational factor provides a good reason to apply game elements to other contexts (Sailer et al. 2017). Specifically, gamification is defined as “the use of game design elements in non-game contexts” (Deterding et al., 2011) (p.9). In the context of education, Sailer and Homner (2020) referred to gamification as gamified learning, which entails “altering existing learning processes to create a revised version of this process that users express as game-like” (p. 78). In gamification, the motivational powers of games are used for other purposes

than entertainment, to enhance learning (Sailer et al., 2013). Additionally, implementing game-like activities into school activities and traditional learning makes them more interesting and attractive as computer games are enjoyable (Attali & Arieli-Attali, 2015). This statement can be supported by a qualitative study on the experiences of gamification in learning performed by Turan et. al (2016). The participants reported that lessons are more enjoyable and interesting due to gamification elements and that this new addition made learning easier.

Landers (2015) has proposed a gamified learning theory through which he tries to understand the relationship between learning and gamification. He explains that a direct mediating process and a moderating process exist by which game elements can affect learning, which form the foundation of his theory. Landers' theory comprises four components, namely instructional content, game characteristics, behaviours and attitudes and learning outcomes, that create these process paths. He mentions that gamification aims to improve instruction rather than replace it, and therefore useful content is a prerequisite for effective gamification. Sailer and Homner (2020) investigated the effects of gamification on several outcomes in the context of education, and found that gamification has positive effects on cognitive, behavioural, and motivational learning outcomes. These findings are in line with Landers' theory of gamified learning. On one hand, it can be concluded that gamification can be used as a mechanism to engage and motivate students.

On the other hand, many studies regarding the topic of gamification in education have been performed that show contrary results. For example, Toda et al. (2017) explored the negative effects of gamification on education. Based on their research and a broad literature search, they identified four main negative effects occurring during quiz-taking: indifference, loss of performance, undesired behaviour, and declining effects. However, there appear to be unexpected side effects too. Andrade et al. (2016) suggest that such negative side effects include undesired competition (i.e. leaderboards could be harmful to students with low performance or low self-efficacy), off-task behaviour (i.e. gamification elements untied to educational outcomes will be distracting to the user), and addiction and dependence (i.e. some elements could be seen as addictive factors, creating a sense of dependency on the game). Thus, it can be concluded that there is still a much on-going debate on whether gamified learning is effective, as there are many positive, as well as negative and unexpected side effects to gamified learning. As Xiao and Hew (2024) have stated, the research on the effectiveness of gamification over the last decade has yielded inconsistent results.

1.3.2 Point Wagering

The types of game elements available vary significantly and therefore can take on many different forms. Sailer et al. (2013) mention eight typical game elements in their article about motivation through gamification, ranging from *quests* to *avatars*, and *badges* to *meaningful stories*. Their first element regards *points*, which is explained as: "Points can be accumulated for certain activities within the gamification environment" (p. 30). One important supporting role of points is the element of feedback to the user (Bai et. al, 2020; Attali & Arieli-Attali, 2015) and the overview of achievements (Bai et. al, 2020). In this research, we explicitly focus on point division by students in tasks. We refer to this as the gamification method of "point wagering". Wagering is defined as being willing to bet money on something because one is convinced it is true (Oxford Education, n.d.). However, in the context of ARSs, the student would not bet money, but assign a certain number of points to the answer(s) that they think is/are correct. This method lies closely with the identification of a student's confidence, which is usually assessed by the question of how confident one is that their chosen answer is the correct answer (Morony et al., 2013). By letting the student 'bet' their points on the answer they believe is true, they show their confidence through their bet on that answer. Simultaneously, this method provides the user with a sense of autonomy and control in their learning.

A range of literature already exists on the topic of points as a gamification element, and has received mixed reviews on its effectiveness. On one hand, problems occurring from points have already been established. For example, Park and Kim (2022) investigated what problems are encountered when this gamification method is applied improperly. They identified two important ways in which points are incorrectly implemented: (1) points that only accumulate weaken the sense of achievement and anticipation, and (2) points emphasize differences between users. On the other hand, however, it has been established that points indeed lead to better learning in students. Hellberg and Moll (2023) investigated this phenomenon. They concluded that points can be beneficial for education, specifically concerning motivation. But, they emphasised that the usage of points is an approach at aiming achievement, and must be implemented together with other pedagogical approaches. This discussion appears to be a part of the above mentioned debate on whether gamified learning is effective as a whole, and therefore acts as a foundation for further investigation on this topic.

1.4 The Effect of Self-Efficacy on Point Wagering

Next to measuring the effectiveness of point wagering, this study also focuses on the effect of academic self-efficacy of the implementation of point wagering. A wide body of scientific research already exists on the effect of gamification on self-efficacy (e.g. Polo-Peña

et al., 2020), and aiming to improve self-efficacy through gamification (e.g. Banfield & Wilkerson, 2014). However, this study aims to investigate the adverse effect, as we will measure the influence of self-efficacy on the approach of using point wagering in quizzes.

Self-efficacy relates to the extent to which students believe they will be able to succeed at a certain task (Ahmadi, 2020). Specifically, according to Bandura (1982), self-efficacy concerns the persistence one shows to overcome obstacles in their way. In an academic context, this refers to the ability of learners to attain their academical goals successfully and has been measured in many learning environments including high schools and universities (Honicke & Broadbent, 2016). Despite the specific learning environments and settings in which academic self-efficacy is measured, a positive correlation exists between academic self-efficacy and academic performance (Honicke & Broadbent, 2016).

Moronoy et al. (2013) make an important distinguishment between self-efficacy and confidence, as they emphasize that 'confidence' is used for both constructs, even though they differ in meaning. They explain that self-efficacy is asked prior to a cognitive task, while confidence is asked after performing that cognitive task. Additionally, Cramer et al. (2009) add that confidence refers to the perception of difficulty of the task, and likelihood that quizzes are provided with the correct answers, while self-efficacy refers to the perception of the ability to perform. Ultimately, the relationship between self-efficacy and test performance has been investigated severely. As Bandura (1997) suggests, self-efficacy is associated with test performance (i.e. a student with higher self-efficacy levels also scores higher on a test). More specifically, academic self-efficacy positively predicts academic performance with regards to grade averages (Salazar & Hayward, 2018; Tilfarlioglu & Ciftci, 2011). Salazar and Hayward (2018) performed a study in which they measured the effect of academic self-efficacy on academic performance and identified a positive correlation between these variables. This finding is in line with Bandura's (1997) statement on self-efficacy.

While considering the link between academic self-efficacy and point wagering, point wagering is a type of gamification element that allows the student to show their confidence in their answers. This implementation of point wagering is best described as a method that regards students as autonomous learners within their learning processes, as the wagering mechanism provides a sense of control. Alkan and Arslan (2019) investigated how autonomous learning behaviours are related to self-efficacy among a range of literature. They concluded that a strong relationship exists between these variables. Additionally, in their study on language potential

and responsibility for a student's role in language learning, Tilfarliogly and Ciftci (2011) also identified a significant positive relationship between self-efficacy and autonomy.

Based on the established relationships between academic self-efficacy and confidence, and academic self-efficacy and autonomy, it becomes interesting to see how self-efficacy affects student behaviour in a quiz with point wagering, which regards these students as autonomous and allows them to demonstrate their confidence.

1.5 Research Question

According to Sailer and Homner (2017), there is little to no analytic evidence on the effectiveness of gamification in the context of education. Therefore, this research will contribute to helping to gain more information on this topic, mainly focusing on the specific gamification element point wagering. Additionally, a wide range of literature exists on the effect of gamification on self-efficacy, but not on the effect of self-efficacy on the utilization and application of gamification elements. Thus, this research will also contribute to establishing the effects of self-efficacy on the utilization of point wagering.

First, this investigation focuses on an alternative variation of the gamification element points. This broadens the widespread body of knowledge that already exists on the effect of gamification elements. This knowledge is mainly focused on the gamification elements that were also provided by Sailer et al. (2013). Second, we will try to close this analytic gap, by specifically focusing on the use of ARSs, instead of other educational tools. This will contribute to an accurate representation of the actual effectiveness of point wagering in ARSs. Lastly, this research contributes to the knowledge existing on the effects of academic self-efficacy on quiz-taking, while simultaneously measuring how this alters if gamification elements are incorporated.

Based on this literature research, the following research questions have been formulated:

(1) RQ1: How does the implementation of the gamification method point wagering affect scores of students while answering quizzes in Audience Response Systems?
(2) RQ2: How does the self-efficacy of students affect their performance in Audience Response Systems in the two answering conditions?
(3) RQ3: How does the self-efficacy of participants affect the way they apply wagering (i.e. show confidence) in Audience Response Systems?

2. Methods

2.1 Participants

To collect data, participants were recruited through the system of SONA of the organisation of the researcher. Additionally, a survey was sent out to friends and family of the researcher, who are students. Exclusion criteria therefore included everyone who was not a student. Methods of sampling used were convenience sampling, snowball sampling, and random sampling. In total, 66 participants responded, of which 18 responses were deleted. Participants were deleted if they did not provide their consent, withdrawn their consent after debriefing, if they did not finish the survey, or were not students. Thus, a sample of 48 participants was analysed, which were randomly divided (by Qualtrics) over a control group and a treatment group. Their demographics are presented in Table 1. Nationalities included Dutch, German, Indian, American, and Estonian. Levels of education included senior secondary vocational education, bachelor's degree, master's degree, and doctoral degrees.

Table 1

Demographics of All Groups

Groups	Full Sample				
	N		M_{age}	*SD*	*Range*
	Male	Female			
Full-Sample	16	32	21.458	1.77	18-25
Control Group	6	20	21.308	1.738	18-25
Treatment Group	10	12	21.636	1.839	18-25

Note. $N = 48$, (control group: $n = 26$, treatment group: $n = 22$).

2.2 Materials

In order to measure the general knowledge of the participants, a knowledge test was created through a survey. To create this survey, Qualtrics was used. The survey consisted of four components: (1) introduction and informed consent, (2) an instrument measuring self-efficacy, (3) a knowledge test (also referred to as a quiz), and (4) a closing statement. A computer, smartphone, or tablet was necessary to participate in this survey.

Introduction. In the introduction, the participants were deceived by an incorrect statement regarding the aim of the survey: "to collect data on student confidence and performance while using Audience Response Systems such as Kahoot, Socrates, and Qualtrics."

After, the participants were informed that this study was completely anonymous and voluntary, and required 15 to 30 minutes to finish. Lastly, they were asked for their informed consent.

Self-efficacy Instrument. For the self-efficacy scale, the College Academic Self-Efficacy Scale (CASES) (Owen & Froman, 1988) was used. This scale consists of 33 questions with five-point Likert scale answer options ranging from 1: *very little* to 5: *quite a lot*. Before listing the questions, the participants were asked: "How much confidence do you have about doing each of the behaviours below? Please select the answer that fits most." Example statements were: "Taking well-organized notes during a lecture," "attending class regularly," or "understanding most ideas presented in class." All statements are listed in Appendix A.

Ultimately, Cronbach's Alpha was performed on the entire self-efficacy instrument. This instrument was evaluated as highly reliable ($\alpha = 0.88$).

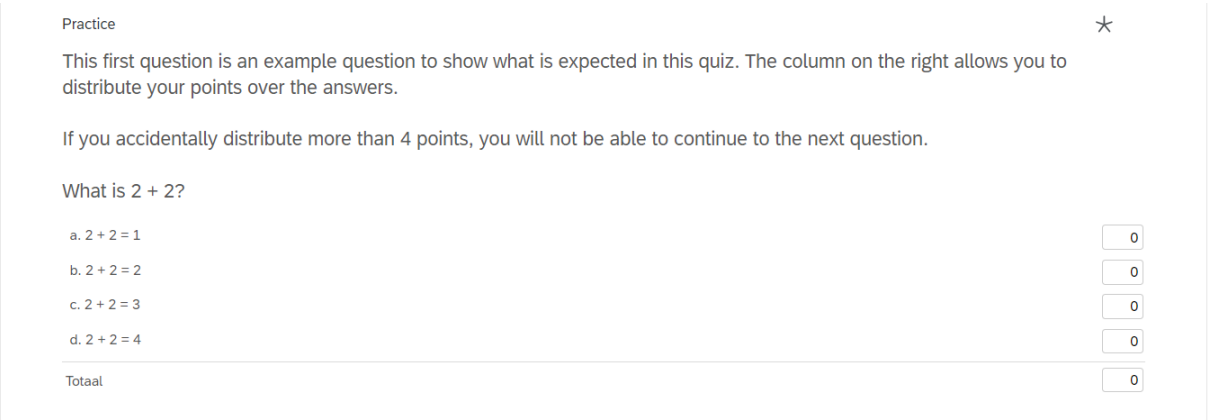
Quiz. To measure the general knowledge of students to see the effects of point wagering and self-efficacy, the following procedure was followed. For the formulation of quiz questions, old Dutch HAVO Cito exams were analysed. Interesting concepts were then fed to an AI generator, which was asked to formulate quiz questions based on these concepts (in the topics of: biology, science, math psychology, geography, and history). Ultimately, these questions were used as guidance, as all questions included in the knowledge quiz have been formulated by the researcher. For the formulation of the answers, the guidelines provided by www.sagaproject.net were followed to ensure properly formulated questions and answer options. These questions were formulated for the six different topics, based on the assumption that in this manner every participant has an area of knowledge in which they excel.

This part of the survey consisted of two versions, both comprised of the same list of questions (Appendix B). Version A included a quiz of 24 multiple-choice questions with four choices each, organized into the six topics. In this version, the participant was given a brief introduction. Ultimately, no example question was shown, and participants immediately proceeded to the first question of the quiz. This introduction is added in Appendix C. Version B included the same 24 questions as Version A but included a wagering mechanism to answer each question. In this version, the participants were given an explicit explanation, along with examples, of how the questions were supposed to be answered. In this explanation, participants were shown how the wagering mechanism worked, through examples of different point variations. An example proposed was: "Option 1. Answer A- 4 points. Answer B- 0 Points. Answer C – 0 points. Answer D – 0 points. *You are certain answer A is the correct answer.*

Therefore, answer A is given all four points." This was repeated for all possible differentiations. The complete explanation with examples is added in Appendix D. Ultimately, these variations measured the 'demonstrated confidence' of a student, which will be explained in detail later in this chapter. After, a simple practice question followed: "What is 2+2?" to experience the wagering mechanism. This example question is shown in Figure 1.

Figure 1

Example Question



Note: Every question was set to a minimum of four points, and a maximum of four points. If answered more points or less points by accident, participants were not able to continue.

Moreover, Cronbach's Alpha was performed on the knowledge quiz. This instrument was evaluated as poor ($\alpha = 0.33$).

Debrief. In the final phase of the survey, participants were informed about the deception in this research. They were informed of the actual aim of this research, which was to measure *the effect of self-efficacy on the gamification method of point wagering and its utilization, as well as measuring the effect of point wagering, while using Audience Response Systems.* Afterward, they were provided with the option to withdraw from the research. In this phase, the participants were not provided with their scores. However, if desired, participants could refer themselves to an answer sheet with detailed explanations of every question through a link.

2.3 Procedure

First, the participants were given a direct link to either SONA (for university students), or to Qualtrics. The survey was distributed through WhatsApp, Instagram, and Snapchat. In addition, the survey was available via the organisations' SONA system, for students that were

not linked to the researcher. In SONA, the study was called *Measuring Confidence and Performance in Audience Response System activities*, with a concise description stating: “This study examines how student confidence affects performance while performing quiz activities in audience response systems”.

After opening, they were introduced to the research. Participants then proceeded with the first part of the survey: filling in the self-efficacy instrument.

Only after finishing the instrument, participants were able to start the second part of the survey: the quiz. Both conditions were not allowed to use a calculator, note sheet, or the internet to look up answers. Additionally, they were told there was no time limit and there was no possibility to return to previous questions. This information was also included in the introductions of both versions.

After finishing the quiz, participants were referred to the closing statement. After closing the survey, their participation ended.

2.4 Data Analysis

For data analysis, RStudio v4.1.1 was used (RStudio, 2020). In R, the packages *tidyverse*, *dyplr*, *psych*, and *broom* were used. For the statistical analysis, $\alpha = 0.05$ was used. Moreover, all test assumptions were checked and non-violated, allowing us to use parametric tests.

2.4.1 Self-efficacy Instrument

For the self-efficacy instrument, the means per participant were calculated. This procedure has been followed for both the control group and the treatment group. Based on the means per participant, the variable “Self-efficacy” was constructed.

2.4.2 Quiz Scores

For the control group, all correct answers were given one point, and incorrect answers were given zero points. Based on this recoded version, the percentage scores of correct scores per participant were calculated (e.g. 15 correct points: 15/24*100 = 62.5%). Ultimately, these scores were constructed into the variable “Quiz Scores”, for the control group.

For the treatment group, all points given to the correct answer were summed up and the percentage scores of correct scores per participant were calculated (e.g. 50 correct points:

50/96*100 = 52.10%. *96 = 4 points * 24*). Ultimately, the scores were constructed into the variable “Quiz Scores”, for the treatment group.

2.4.3 T-Test

A two-sample t-test has been performed on the variables “Quiz Scores” and “Self-efficacy” to compare the means of the control group and treatment group.

2.4.4 Analysis of ‘Demonstrated Confidence’ and Analysis of Different Patterns Used

Different variations of point distributions used in the quiz of the treatment condition were analysed, evaluated, and recoded. To specify, all five possible different variations received a different number/code. This number/code simultaneously represents the amount of demonstrated confidence shown in the wagering quiz.

Variation (5) - (5 points): Variation 4-0-0-0 (in any order)

Variation (4) - (4 points): Variation 3-1-0-0 (in any order)

Variation (3) - (3 points): Variation 2-2-0-0 (in any order)

Variation (2) - (2 points): Variation 2-1-1-0 (in any order)

Variation (1) - (1 point): Variation 1-1-1-1 (in any order)

This rule applies because of the hypothesis that the more confident a student is in their answer, the less they will distribute their points over different answers. Similarly, this was explained to the participants during the introduction of the wagering quiz (see Appendix C).

Based on this recoding, the means per participant were calculated. These scores were constructed into a new variable, namely “Demonstrated Confidence”, representing the *demonstrated confidence shown in the wagering quiz*, in which 1.00 equals the lowest possible demonstrated confidence, and 5.00 equals the highest possible demonstrated confidence.

Lastly, it was calculated how often a different variation was used during the wagering quiz.

2.4.5 Correlation Analysis

Pearson’s bivariate correlation coefficient was used to compute correlations between the variables “Quiz Scores”, “Self-efficacy” and “Demonstrated Confidence”.

3. Results

3.1 Descriptives

For all variables, the mean of the quiz scores, standard error of the mean, and standard deviation of the mean were calculated. The results are shown in Table 2.

Table 2

Descriptive Statistics

Scale	Group	*n*	*M*	*SD*	*SE*
Quiz Scores	C	26	53.045	10.964	2.150
	T	22	57.434	14.247	3.038
Demonstrated	C	0	-	-	-
Confidence	T	22	4.282	0.479	0.102
Self-efficacy	C	26	3.599	0.490	0.096
	T	22	3.488	0.463	0.099

Note. C = control group, T = treatment group.

3.2 T-Test

A two-sample t-test has been performed on the variables “Quiz Scores” and “Self-efficacy”, shown in the table above. For both variables, the analysis showed no significant differences between the groups (Quiz scores: $t(46) = 1.210$, $p = .234$) (Self-efficacy: $t(46) = 0.544$, $p = .589$).

3.3 Correlation Analysis

Pearson’s correlation analysis has been performed on the following variables:

(1) “Quiz Scores” and “Self-efficacy” for the control condition.

(2) “Quiz Scores”, “Demonstrated Confidence” and “Self-efficacy” for the treatment condition.

For the control group, Pearson’s bivariate correlation test showed no significant correlation between the variables “Quiz Scores” and “Self-efficacy” ($r[24] = -.03$, $n = 26$, $p = .872$). For the treatment group, Pearson’s bivariate correlation test showed a significant positive correlation between the variables “Quiz Scores” and “Demonstrated Confidence” ($r[20] = .54$, $n = 22$, $p = .010$), and between the variables “Quiz Scores” and “Self-efficacy” ($r[20] = .43$, $n = 22$, $p = .046$). All Pearson’s correlations found in the treatment group are reported in Table 3.

Table 3

Correlations for Study Variables "Quiz Scores", "Demonstrated Confidence" and "Self-efficacy" for the treatment condition

Variable	2	3
1. Quiz Scores	.54*	.43*
2. Demonstrated Confidence	-	.05
3. Self-efficacy		-

Note. $N = 22$. *p <0.05.

3.4 Pattern Analysis

Table 4 shows how often variations 5, 4, 3, 2, and 1 (most to least points) have been used during the quiz by participants in the treatment condition.

Table 4

Frequency Table of Different Variations Used in Treatment Condition

Variation/code	Variation specified (in any order)	Frequency	Percentage
(5)	4-0-0-0	381	72.16%
(4)	3-1-0-0	27	4.92%
(3)	2-2-0-0	57	11.17%
(2)	2-1-1-0	14	2.65%
(1)	1-1-1-1	49	9.09%

Moreover, it became evident that the treatment condition assigned the most points to the correct answer in 20 out of 24 questions. This was not true for questions 8, 16, 20 and 24, indicating difficult or unclear items.

4. Discussion

A wide range of research has been performed on the topic of gamification in the context of education and learner performance. However, there is a clear debate about the significance of its effect. The current study performed an analysis comparing the results of students who were exposed to the gamification element of point wagering in a knowledge tes to the results of students who were not exposed to this gamification element. In addition, the influence of student's self-efficacy in an academic context on performance was measured, as well as the ways in which wagering was applied. The three research questions remain as follows:

(1) RQ1: How does the implementation of the gamification method point wagering affect scores of students while answering quizzes in Audience Response Systems?
(2) RQ2: How does the self-efficacy of students affect their performance in Audience Response Systems in the two answering conditions?
(3) RQ3: How does the self-efficacy of participants affect the way they apply wagering (i.e. show confidence) in Audience Response Systems?

4.1 RQ1: How does the implementation of the gamification method point wagering affect scores of students while answering quizzes in Audience response systems?

After analysing the data, it became evident that there is no significant difference between quiz scores of the control group and the treatment group, as the t-test showed non-significant results. Therefore, it can be concluded that there was no effect of the implementation of the wagering mechanism on the quiz scores.

The findings of this study relate to the findings of other literature to some extent. A review performed by Manzano-León et. al (2021), aimed at gaining a clearer understanding of gamification in education, explored the effect of multiple gamification elements, including points. Their research revealed a different finding than our study, as they discovered that points made it possible for their participants to improve their overall scores. Accordingly, they found a significant correlation between gamified activity and quiz scores. These results clearly contradict the findings of our study. Interestingly, Attali and Arieli-Attali (2015) have also investigated the effect of gamification on performance. In contrary to the former study, the findings of Attali and Arieli-Attali (2015) confirm our findings that the gamification element of points has no effect on its own if self-efficacy is not considered. In their study, they investigated the influence of points on performance in computerized assessments of basic

mathematical concepts. Their findings concluded that the gamification element did not facilitate the accuracy of responses in both conditions.

These findings further contribute to the ongoing debate on whether points are an effective gamification method. As illustrated by Xiao and Hew (2024), inconsistent results to the research on its effectiveness have been found that are either significant, non-significant, or even have a negative impact on participants. The non-significant difference in quiz sores can be explained by different factors. First, some contextual factors could contribute to the ineffective implementation of points, namely the subject matter and the instructor (Rabah et al. 2018). As could be concluded by the pattern analysis, and average scores of the quiz, the quiz could be evaluated as difficult. Ultimately, without the presence of an instructor, and rather taking the quiz alone, a combination of these factors could possibly also affect the utilization of point wagering of the students. Moreover, Hellberg and Moll (2023) stated that the gamification element of points is only successful alongside other pedagogical approaches. However, as this quiz was taken outside student curriculums, and was based on a knowledge test rather than one subject topic, this could also have resulted in poorer results. Lastly, another reason for our findings could be the so-called ‘one-size-fits-all’ approach which potentially mitigates the effects of points, as it does not tailor the quiz to the user’s attributes (Rodrigues et al., 2023). This suggests that for the gamification elements to be successful in ARSs, one should consider the preferences of students for specific gamification types, instead of assuming one type of gamification.

4.2 RQ2: How does the self-efficacy of students affect their performance in Audience Response Systems in the two answering conditions?

For the control group, Pearson's bivariate correlational analysis showed there is no significant correlation found between variables ‘Quiz Scores’ and ‘Self-efficacy’. But, for the treatment group, the correlation between these variables is positively significant.

Our results showed that in a standard multiple-choice quiz lay-out, self-efficacy did not influence the quiz scores of students. This differs from the findings of many other studies, that explained how a higher self-efficacy also means higher scores on exams (Feldman et al, 2011; Barrows et al., 2013). However, it is apparent that self-efficacy is not the only indicator of academic performance (e.g. capability, student habits, intelligence, and motivation) (Barrows et al., 2013; Salazar & Hayward; 2018). Chou (2019) examined what elements contribute to self-efficacy in test taking. Their results showed that the best predictor of self-

efficacy was either test anxiety or preparation. Considering the latter, students that were well-prepared for the exam by partaking in all preparations, believed they would perform better. This is an important observation, as our study did not give students an opportunity to prepare for their exam, and were rather given the test without any further information on the content (except for the different topics). Accordingly, our quiz aimed to measure different topics so that all participants would be able to at least answer one topic of knowledge. Considering the inability to prepare for this test, as well as knowledge on the topic (referring to capability and intelligence), this could have had a negative impact on self-efficacy.

Oppositely, Chou (2019) also mentioned the effect of test anxiety on self-efficacy. Barrows et al. (2013) mention that empirical evidence exists on the relationship between academic performance, self-efficacy, and test anxiety. More specifically, high self-efficacy is related to higher academic performance, while high test-anxiety is related to lower academic performance. Therefore, a possibility exists that other personal characteristics affected self-efficacy and performance, but these elements were not considered in this research.

4.3 RQ3: How does the self-efficacy of participants affect the way they apply wagering (i.e. show confidence) in Audience Response Systems?

Pearson's bivariate correlation analysis showed that the correlation between the variables 'Quiz Scores' and 'Demonstrated Confidence' is positively significant for the treatment group. Interestingly, there is no significant correlation found between variables 'Self-Efficacy' and 'Demonstrated Confidence' in this group. Based on these findings it appears that participants with higher self-efficacy will also score higher on a quiz and have a higher demonstrated confidence.

Moreover, the results of the pattern analysis show that, even though participants in the treatment condition fully understood that different variations could be applied in the wagering answering options, participants used the fifth variation (4-0-0-0 (in any order)) almost 75% of the time. Thus, participants will, in many cases, go for the point distribution in which they show most confidence. In addition to this observation, we saw mean scores of 53.05% and 57.43% in both groups for the quiz scores. This indicates that the quiz was rather difficult, as only half of the points were obtained, and it can be concluded that students will not pass the quiz in many cases. Considering this finding for the treatment group, many students demonstrated the highest amount of confidence in almost 75% of the amount of variations

asked, but obtained an average score of 57%, indicating that their confidence is often fully put in the wrong answer.

This interesting finding could be explained through the phenomenon of 'calibration': the relationship between one's actual performance and their metacognitive judgements about it (Papadopoulos et al. 2021a), in which the students in the treatment condition showed low calibration. Dunlosky and Rawson (2012) attempt to explain this phenomenon. In their study, they investigated the effect of overconfidence, and concluded that a strong relationship exists between one's judgement accuracy and long-term retention. They showed that students that were not over-confident showed better performance during practice sessions and therefore led to better learning, in comparison to over-confident students that performed poorer. Consequently, León, Lipnevich and Ferrero (2024) add that a positive link between calibration and achievement has already been identified, in which students that are well-calibrated set more realistic goals and are more able to determine what needs to be done to improve their performance.

Moreover, our research showed that students with a higher level of self-efficacy are more likely to use higher variations in the wagering mechanism, and therefore demonstrate a higher confidence in their answers. This finding relates to statement of Morony et al., (2013) explaining confidence, which involves that confidence is usually assessed by how confident a student is that their chosen answer is also the correct answer. It can therefore also be confirmed that self-efficacy and confidence are closely linked to each other, as previously established (Cramer et al., 2009). Therefore, self-efficacy is an indicator of the demonstrated confidence shown by students, as students with higher self-efficacy apply higher confidence patterns. When students are allowed more autonomy and control in their answers through point wagering, self-efficacy influences their quiz scores.

4.4 Limitations

An important limitation of this research is that this quiz was taken online, instead of in a classroom. Firstly, in a scenario where both conditions would take their quiz in a classroom with a teacher, there would be control for Internet use, use of calculators, and taking the quiz with others during the quiz. However, since participants took their quiz on their own, this could not be controlled. Unfortunately, it remains unclear if participants adhered fully to the rules of the quiz in their condition, which affects the validity of this research. Ultimately, a chance exists that these rules were not adhered to, and therefore could affect the overall scores of the quizzes.

Second, there remains a possibility that participants in the treatment condition had questions about what was expected of them. For these students, a teacher would also be of help.

Another limitation concerns the use of this specific quiz instrument. The knowledge test was specifically constructed for this research by the researcher and thus has not been used before. Due to time constraints, it was not possible to perform a pilot test and evaluate which items were too easy or too difficult in comparison to other items. This could be a possible cause of the lower reliability of the instrument. Moreover, an important factor of both the element of points and of ARSs has been overlooked: feedback to the student during the quiz. As mentioned in the introduction, students learn because they receive feedback on their performance. However, since this feedback element was missing, students did not see their scores after each question.

4.5 Conclusions and Future Research

During our literature search, it became evident that an analytic gap exists in the knowledge on the effect of gamification elements in education. In this specific research, we narrowed down the scope to ARSs (i.e. quiz activities), so that a more targeted examination of the impact of gamification was provided, in this specific learning context. A widespread of literature already exists on the use of gamification in educational tools, specifically *points*. This research introduced a new version of this gamification element, namely point wagering, to measure its effectiveness. Simultaneously, this study also measured how self-efficacy affects the performance and utilization of point wagering, and therefore expands the knowledge by approaching the topic of gamification in education from a different viewpoint. The results show no significant differences in the quiz scores of both conditions. Additionally, no effect of self-efficacy on the quiz scores was found for students in the control group. However, a positive correlation was found between the quiz scores and self-efficacy of students in the treatment group, as well as a positive correlation between their quiz scores and demonstrated confidence. This suggests that in a quiz that allows for control by students, self-efficacy affects the level of confidence shown by students, which ultimately affects their quiz scores.

4.5.1 Future Research

Based on the knowledge gained during this research, suggestions could be made for future research. First, the effect of self-efficacy on performance and the utilization of point wagering should be measured in an onsite scenario, where the rules provided for the quiz (no Internet use, no calculator or note sheets, and no discussion among peers) can be controlled.

Creating an onsite test environment would tackle the problem of the validity of our research. The second suggestion considers the reliability of the quiz instrument. To improve the reliability of the quiz, a pilot test must be performed, to indicate which items did not fit in the instrument. The pilot test would help improve the quality of the items included in the instrument, and check if items where included that were found to be too difficult or too easy. In addition, the quiz should be adjusted in such a way that a feedback element is included. As this was an important overlooked factor, this element should be included in future research, as feedback is an important element of both ARSs and the gamification element of points. The last suggestion includes the decision for a second test moment. The results provided by Marin et al. (2018) and Manzano-León (2021) showed different effects of gamification elements on student's learning progress. By integrating a second test moment for both conditions, the effects of point wagering could be measured after a period (e.g. midterm and end term). Based on that information, the conclusion could be drawn whether point wagering improves grades during a semester or if student's quiz scores remain constant.

4.5.2 Implications

Our findings are useful for practice for teachers, or other instructors in the educational fields. First, as this research provides a clear view of how self-efficacy affects the utilization of point wagering, this knowledge could help teachers focus on and support students with low academic self-efficacy levels by providing them with useful strategies or tips. Moreover, as it became evident that point wagering does not improve quiz scores on its own, the focus lies on self-efficacy to enhance the learning outcome. Second, the findings of this research could help businesses and organisations create learning tools that are effective in enhancing self-efficacy, as well as enhancing the learning progress of students. This could be achieved by creating a tool that corrects the issues related to the difficulty of items and implements feedback.

4.5.3 Contributions

Revisiting the contributions of this study, this study has succeeded in gaining more knowledge on the topic of the specific gamification element of point wagering in the context of audience response systems. Moreover, it has succeeded at establishing the effect of self-efficacy on quiz-taking in a gamified environment that provides control to the user. Therefore, it attempted to close the analytic gap on the effectiveness of gamification in education (Sailer & Homner, 2017), attempted to close knowledge gap on the effect of self-efficacy on point

wagering, and attempted to narrow this scope to focus on the effect of self-efficacy in this context.

Sources

Ahmadi, S. (2020). Academic self-esteem, academic self-efficacy and academic achievement: A path analysis. *Journal of Forensic Psychology, 5*(1). https://doi.org/10.35248/2475-319X.19.5.155

Alkan, M.F., & Arslan, M. (2019). Learner autonomy of pre-service teachers and its associations with academic motivation and self-efficacy. *Malaysian Journal of Learning and Instruction, 16*(2), 75-96.

Andrade, F.R.H., Mizoguchi, R., & Isotani, S. (2016). The bright and dark side of gamification. In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Intelligent tutoring systems. ITS 2016. Lecture notes in computer science: Vol: 9684.* Springer. https://doi.org/10.1007/978-3-319-39583-8_17

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: Do points affect test performance? *Computers & Education, 83*, 57-63. https://dx.doi.org/10.1016/j.compedu.2014.12.012

Bai, S., Hew, F.K., & Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review, 30*. https://doi.org/10.1016/j.edurev.2020.100322

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147. https://doi.org/10.1037/0003-066X.37.2.122

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman & Company.

Banfield, J., & Wilkerson, B. (2014). Increasing student intrinsic motivation and self-efficacy through gamification pedagogy. *Contemporary Issues in Education Research – Fourth Quarter 2014, 7*(4), 291-298.

Barrows, J., Dunn, S., & Lloyd, C.A. (2013). Anxiety, self-efficacy, and college exam grades. *Universal Journal of Educational Research, 1*(3), 204-208. https://doi.org/10.13189/ujer.2013.010310

Chou, M. (2019). Predicting self-efficacy in test preparation: Gender, value, anxiety, test performance and strategies. *The Journal of Educational Research, 112*(1), 61-71. https://doi.org/10.1080/00220671.2018.1437530

Cramer, R.J., Neal, T.M.S., & Brodsky, S.L. (2009). Self-efficacy and confidence: Theoretical distinctions and implications for trial consultation. *Consulting Psychology Journal: Practice and Research, 61*, 319-334. https:///doi.org/10.1037/a0017310

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining "gamification". In A. Lugmayr (Ed.), Proceedings of the 15th International Academic Mindtrek Conference: Envisioning Future Media Environments (pp. 9–15). New York: ACM. https://doi.org/10.1145/2181037.2181040

Donohue, S. (2014). Supporting active learning in an undergraduate geotechnical engineering course using group-based audience response systems quizzes. *European Journal of Engineering Education, 39*(1), 45-54. https://doi.org/10.1080/03043797.2013.833169

Dunlosky, J. & Rawson, K.A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine student's learning and retention. *Learning and Instruction, 22*, 271-280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Edmonds, C.T, & Edmonds, T.P. (2008). An empirical investigation of the effects of SRS technology on introductory managerial accounting students. *Issues in Accounting Education, 23*(30), 421-434.

Feldman, E., Kim, J.S., & Elliot, S.N. (2011). The effects of accommodations on adolescents' self-efficacy and test performance. *The Journal of Special Education, 45*(2), 77-88. https://doi.org/10.1177/0022466909353791

Gleason, B.L., Peeters, M.J., Resman-Targoff, B.H., Karr, S., McBane, S., Kelley, K., Thomas, T., & Denetclaw, T.W. (2011). An active-learning strategies primer for achieving ability-based educational outcomes. *American Journal of Pharmaceutical Education, 75*(9).

Hellberg, A., & Moll, J. (2023). A point with pointsification? Clarifying and separating pointisfication from gamification in education. *Front. Educ., 8*. https://doi.org/10.3389/feduc.2023.1212994

Honicke, T., & Broadbent, J. (2016). The relation of academic self-efficacy to university student academic performance: A systematic review. *Educational Research Review, 17*, 63-84. https://doi.org/10.1016/j.edurev.2015.11.00

Hunsu, N.J., Adesope, O., & Bayly, D.J. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Computers & Education, 94*, 102-119. https://doi.org/10.1016/j.compedu.2015.11.013

Kay, R.H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education 53*(3), 819-827. https://doi.org/10.1016/j.compedu.2009.05.001

Landers, R.N. (2015). Developing theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6). https://doi.org/10.1177/1046878114563660

León, S. P., Lipnevich, A.A., & Ferrero, M. (2024) Differences is student's performance, (mis)calibration, and confidence judgments depending on their preferred sitting position in class. *Personality and Individual Differences, 217*. https://doi.org/10.1016/j.paid.2023.112424

Mader. S., & Bry. F. (2019). Audience response systems reimagined. In M. Herzorg, Z. Kubincová, P. Han, & M. Temperini (Eds.), *Advances in web-based learning – ICWL 2019. Lecture notes in computer science: Vol: 18841* (pp. 203-216). https://doi.org/10.1007/978-3-030-35758-0_19

Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M.A., Guerrero-Puerta, L., Aguilar-Parra, J.M., Trigueros, R., & Alias, A. (2021). Between level up and game over: A systematic literature review of gamification in education. *Sustainability, 13*. https://doi.org/10.3390/su13042247

Marin, B., Frez, J., Cruz-Lemus, J., & Genero, M. (2018). An empirical investigation on the benefits of gamification in programming courses. *ACM Transactions on Computing Education, 19*(1). https://doi.org/10.1145/3231709

Mazur, E. (2009). Farewell, lecture? *Science, 323,* 50-51.

Morony, S., Kleitman, S., Lee, Y.P., & Stankov, L. (2013). Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research, 58*, 79-96. https://doi.org/10.1016/j.ijer.2012.11.002

Ng, W. (2015). *New digital technology in education*. Springer. https://doi.org/10.1007/978-3-319-05822-1

Nicholson, B.T., & Bassaignani, M.J. (2009) Radiologist/educator knowledge of the audience response system and limitations to its use. *Academic Radiology, 16*(12), 1555-1560. https://doi.org/10.1016/j.acra.2009.07.014

Owen, S. V., & Froman, R. D. (1988). Development of a college academic self-efficacy scale.

Oxford Dictionary. (n.d.) Wagering. In *Oxford Dictionary.* Retrieved on February 26, 2024, from https://www.oxfordlearnersdictionaries.com/definition/english/wager_2?q=wagering

Papadopoulos, P.M., Obwegeser, N., & Weinberger, A. (2021). Concurrent and retrospective metacognitive judgements as feedback in audience response systems: Impact on performance and self-assessment accuracy. *Computers and Education Open, 2*. https://doi.org/10.1016/j.caeo.2021.100046

Papadopoulos, P.M., Obwegeser, N., & Weinberger, A. (2021). Let me explain! The effects of writing and reading short justifications on student's performance, confidence and opinions in audience response systems. *Journal of Computer Assisted Learning, 38*(2), 327-337. https://doi.org/10.1111/jcal.12608

Park, S., & Kim, S. (2022). Points and the delivery of gameful experiences in a gamified environment: Framework development and case analysis. *JMIR Serious Games*, 10(3), https://doi.org/10.2196/35907

Polo-Peña, A.I., Frías-Jamilena, D.M., & Fernández-Ruano, M.L. (2020). Influence of gamification on perceived self-efficacy: gender and age moderator effect. *International Journal of Sports Marketing and Sponsorship, 22*(3), 453-476. https://doi.org/10.1108/IJSMS-02-2020-0020

Pradhan, A., Sparano, D., & Ananth, C.V. (2005). The influence of an audience response system on knowledge retention: An application to resident education. *American Journal of Obstetrics and Gynaecology, 195*(5), 1827-1830. https://doi.org/10.1016/j.ajog.2005.07.075

Rabah, J., Cassidy, R., & Beauchemin, R. (2018). Gamification in education: Real benefits or edutainment. *European Conference on eLearning.* https://doi.org/10.13140/RG.2.2.28673.56162

Rodrigues, L., Palomino, P.T., Toda, A.M., Clock, A.C.T., Pessoa, M., Pereira, F.P., Oliveira, E.H.T, Oliveira, D.F, Cristea, A.I., Gasparini. I., & Isotani, S. (2024). How personalization affects motivation in gamified review assessments. *International Journal of Artificial Intelligence in Education, 34*, 147-184. https://doi.org/10.1007/s40593-022-00326-x

RStudio Team (2020). *RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. https://www.rstudio.com/*

Ruhl, K.L, Hughes, C.A., & Schloss, P.J. (1987). Using the pause procedure to enhance lecture recall. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Childre 10*(1). https://doi.org/10.1177/088840648701000

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*, 77-112. https://doi.org/10.1007/s10648-019-09498-w

Sailer, M., Hense, J., Mandl, H., & Klevers, M. (2013). Psychological perspectives on motivation through gamification. *Interaction Design and Architecture(s) Journal, 19*, 28-37. https://doi.org/10.55612/s-5002-019-002

Sailer, M., Hense, U.J., Mayr, S.K., & Mandl, H. (2017) How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69,* 371-380. https://doi.org/10.1016/j.chb.2016.12.033

Salazar, L.R., & Hayward, S.L. (2018). An examination of college students' problem-solving self-efficacy, academic self-efficacy, motivation test performance and expected grade in introductory-level economics courses. *Decision Sciences Journal of Innovative Education, 16*(3).

Schmidt, B. (2011). Teaching engineering dynamics by use of peer instruction supported by an audience response system. *European Journal of Engineering Education, 36*(5), 413-423. https://doi.org/10.1080/03043797.2011.602185

Senthamarai, S. (2018). Interactive teaching strategies. *Journal of Applied and Advanced Research, 3*(1). https://doi.org/10.21839/jaar.2018.v3S1.166

Silliman, S.E., & McWilliams, L. (2004). Observations on benefits/limitations of an audience response systems. In E. Matson, & S. Deloach (Eds.), *Proceedings of the 2004 American society for engineering education annual conference and exposition,* (pp. 1 -16).

Thomas, C.M., Monturo, C., & Conroy, K. (2011). Experiences of faculty and students using an audience response system in the classroom. *Comput Inform Nurs., 29*(7). https://doi.org/10.1097/NCN.0b013e3181fc405b

Tilfarlioglu, F.Y., & Ciftci, F.S. (2011). Supporting self-efficacy and learner autonomy in relation to academic success in EFL classrooms (a case study). *Theory and Practice in Language Studies, 1*(10) 1284-1294. https://doi.org/10.4304/tpls.1.10.1284-1294

Toda, A. M., Valle. P.H.D., & Isotani, S. (2017). The dark side of gamification: An overview of negative effects of gamification in education. In A.I. Cristea, I.G. Bittencourt, & F. Lima (Eds.), *Higher education for all. From challenges to Novel Technology-Enhanced Solutions* (pp. 143–156). https://doi.org/10.1007/978-3-319-97934-2

Turan, Z., Avinc, Z., Kara, K., & Goktas, Y. (2016). Gamification and education: Achievements, cognitive loads, and views of students. *International Journal of Emerging Technologies in Learning, 11*(7). https://doi.org/10.3991/ijet.v11i07.5455

Wood, R., & Shirazi, S. (2020). A systematic review of audience response systems for teaching and learning in higher education: The student experience. *Computers & Education, 153* https://doi.org/10.1016/j.compedu.2020.103896

Xiao, Y. & Hew, K.F. (2024). Personalized gamification versus one-size-fits-all gamification in fully online learning: Effects on student motivational, behavioral and cognitive outcomes. *Learning and Individual Differences, 113*. https://doi.org/10.1016/j.lindif.2024.102470

Appendix A

All correct answers are underlined.

Biology

1. ***What is the main function of red blood cells?***
    a. Red blood cells help the body fight against infections and diseases.
    b. Red blood cells help prevent unwanted substances from entering the body.
    c. Red blood cells transport oxygen to tissue.
    d. Red blood cells allow glucose in the blood to enter other cells, which provides them with energy to function.
2. ***The body contains many diverse types of protein. One of these proteins provides support, elasticity and strength to connective tissues, the skin, and bones, and helps with wound healing. What is the name of this protein?***
    a. Haemoglobin
    b. Keratin
    c. Collagen
    d. Melanin
3. ***Which of the following functions is a primary function of the respiratory system?***
    a. The digestion of foods and absorbing nutrients.
    b. The pumping of blood throughout the body.
    c. The regulation of body temperature.
    d. The exchange of gases with the environment.
4. ***A specialized pigment is found in the eyes of nocturnal animals. This pigment enhances their ability to see in conditions with little light. What is the name of this pigment?***
    a. Rhodopsin
    b. Retinol
    c. Chlorophyll
    d. Carotenoids

Science

5. ***For what appraised work is Galileo Galilei most famous for?***
    a. His telescopic observations and the principle of inertia.
    b. His discovery of the Laws of Motion and Universal gravitation.
    c. His discovery of penicillin.

d. His theories of relativity and his other contributions to physics.

6. Which statement best describes the phenomenon of diffraction?

a. When a sound or light wave strikes an object, it vibrates at a particular frequency. If this frequency matches the resonant frequency of the object it hits, diffraction occurs.

b. Diffraction occurs when several waves are added together, provided that the phase differences between them remain constant over the observation time.

c. Diffraction entails the increase of decrease in the frequency of a wave (sound, light or other) in relation to an observer who is moving relative to the source of wave.

d. Diffraction is the process of light bending around an obstacle or spreading out after moving through small spaces.

7. Why do objects sink or float in fluids?

a. It depends on the density of the fluid.

b. It depends on the density of both the object and fluid.

c. It depends on the weight of an object.

d. It depends on both weight and volume of an object.

8. What is the SI unit of electric charge?

a. Ampere (A)

b. Joule (J)

c. Coulomb (C)

d. Volt (V)

Math

9. 3+3(3*3) =

a. 21

b. 30

c. 34

d. 36

10. What number is not a prime number?

a. 47

b. 79

c. 87

d. 97

11. A farmer owns a piece of land with width 35 and length 60 meters. He plans to expand his land with 25 percent. What will the area of the new land be?

a. 2100 m^2
b. 2520 m^2
c. 2625 m^2
d. 2725 m^2

12. According to the Pythagorean theorem, if side A of a triangle equals 5 cm, and side C equals 13 cm, what will the length of side B be?

a. 8 cm
b. 12 cm
c. 64 cm
d. 144 cm

Psychology

13. In which mental disorder does a patient experience periods of depression and periods of elevated moods and mania?

a. Generalized anxiety disorder
b. Bipolar disorder
c. Schizophrenia
d. Obsessive compulsive disorder

14. According to Jean Piaget's theory of cognitive development, in what stage does the child begin to use language and engages in symbolic play?

a. Stage 1: sensorimotor intelligence
b. Stage 2: preoperational thinking
c. Stage 3: Concrete operational thinking
d. Stage 4: Formal operational thinking

15. In which type of bias does the "I knew it all along" phenomenon occur?

a. Hindsight bias
b. Self-serving bias
c. Confirmation bias
d. Recall bias

16. What are the main characteristics of an authoritative parenting style?

a. Unsupportive parenting style, with no limits to monitoring behaviour. The parents are low in demandingness and sometimes even neglectful.
b. Cold and unresponsive parenting style, in which parents are high in control and demandingness. In addition, parents often make use of parental power.

c. Responsive parenting style, but parents are overly lenient. In addition, there is no requirement for children to regulate themselves.
d. Warm and responsive parenting style, with clear demands, standards, and limits.

Geography

17. What is the process called by which broken down material is transported from one location to another by wind, water, ice, or gravity

a. Weathering.
b. Erosion.
c. Deposition.
d. Sedimentation.

18. Our climate is changing rapidly, due to global warming. Considering coastal regions, which of the following statements is a predicted consequence of climate change?

a. Rising sea levels
b. Decreased temperatures
c. Cooling down of ice caps
d. Increased levels of precipitation

19. The Ozone layer serves as a protective layer around the Earth. What is not one of its roles?

a. The ozone layer contributes to the energy balance of Earth, by absorbing visible light.
b. It fights global warming. Its removal would cause heat to be trapped within the atmosphere.
c. The ozone layer acts as a barrier against harmful gases from outside the atmosphere.
d. The ozone layer acts as a protective barrier that helps prevent UV radiation reach Earth.

20. After eruption, volcanos pose different hazards to aviation. Which of the following poses a risk due to dispersion of volcanic ash particles into the atmosphere?

a. Lahar
b. Pyroclastic flow
c. Ashfall
d. Volcanic bomb

History

21. Which ancient civilization is credited for developing Cuneiform, the earliest known writing system?

a. Ancient Egypt

b. Ancient Mesopotamia

c. Ancient Greece

d. Ancient China

22. Which of the following events marked the beginning of the Age of Exploration in the 15th century?

a. The circumnavigation of the globe by Ferdinand Magellan

b. The establishment of the Cape Route to India by Vasco da Gama

c. The discovery of the New World by Christopher Columbus

d. The colonization of Brazil by the Portuguese

23. The Neolithic Age, or New Stone Age (10,000 years BCE – 3500 BCE), saw significant developments: transitioning from hunting to agriculture, the start of settled communities, and the domestication of animals. What impact did the transition from hunting to agriculture have on the population?

a. Rise of nomadic lifestyles

b. Poor division of labour

c. Decrease of population growth

d. Increase of population growth

24. The Renaissance (14th to 16th century) is marked as a period in history which had a strong fixation on the classical antiquity, in different aspects of life (e.g. music, literature and architecture). Which intellectual movement was a characteristic of this period?

a. Scholasticism

b. Humanism

c. Absolutism

d. Empiricism

Appendix B

Self-efficacy Scale (CASES) Items

(1) Taking well-organized notes during a lecture
(2) participating in a class discussion
(3) Answering a question in a large class
(4) Answering a question in a small class
(5) Taking “objective” tests (multiple-choice, true/false, matching)
(6) Taking essay tests
(7) Writing a high-quality term paper
(8) Listening carefully during a lecture on a difficult topic
(9) Tutoring another student
(10) Explaining a concept to another student
(11) Asking a professor in class to review a concept you do not understand
(12) Earning good marks in most courses
(13) Studying enough to understand the content thoroughly
(14) Running for student government office
(15) Participating in extracurricular events (sports, clubs)
(16) Making professors respect you
(17) Attending class regularly
(18) Attending class consistently in a dull course
(19) Making a professor think you are paying attention in class
(20) Understanding most ideas you read in your texts
(21) Understanding most ideas presented in class
(22) Performing simple math computations
(23) Using a computer
(24) Mastering most content in a math course
(25) Talking to your professor privately to get to know them
(26) Relating course content to material in other courses
(27) Challenging a professor’s opinion in class
(28) Applying lecture content to a laboratory session
(29) Making good use of the library
(30) Getting good grades
(31) Spreading out studying instead of cramming

(32) Understanding difficult passages in textbooks

(33) Mastering content in a course you are not interested in

Appendix C

Qualtrics Explanation Quiz Control Group

Now, you will start the quiz.

This quiz consists of six topics: biology, science, math, psychology, geography, and history. Each topic consists of four questions. All questions contain four multiple-choice answers. Please select the one you believe fits most. You are not able to go back to previous questions.

During the quiz, you are not allowed to use a note sheet, calculator or look up answers on the internet.
This quiz has no time limit.

As mentioned, this research is entirely anonymous. Therefore, read and follow the instructions carefully, because no data can and will be traced back to the participant. But, most importantly, have fun!

Good luck!

Appendix D

Qualtrics Explanation Quiz Treatment Group

Now, you will start the quiz. This quiz consists of six topics: biology, science, math, psychology, geography, and history. Each topic consists of four questions.

All questions contain four multiple-choice answers. In this exam, you select the right answer by assigning points to the answer that you believe is true. You are given a total of four points, which you can assign to the answer options in different ways. Below, options and explanations are provided to show what is expected.

The following are examples of how the multiple-choice questions could be answered. After this explanation, a simple practice question follows in which you can try out the wagering mechanism.

Option 1
Answer A - 4 points
Answer B - 0 points
Answer C - 0 points
Answer D - 0 points

You are certain answer A is the correct answer. Therefore, answer A is given all four points.

Option 2
Answer A - 2 points
Answer B - 2 points
Answer C - 0 points
Answer D - 0 points

You are certain that answer C and answer D are false. However, you are hesitating between answer A and answer B. Therefore, you give two points to answer A and two points to answer B.

Option 3

Answer A - 3 points
Answer B - 1 point
Answer C - 0 points
Answer D - 0 points

You are certain that answer C and answer D are false. However, you are in doubt between answer A and answer B. You believe answer A is more likely to be correct than answer B, and therefore give answer A three points and answer B one point.

Option 4
Answer A - 2 points
Answer B - 1 point
Answer C - 1 point
Answer D - 0 points

You are certain that answer D is false. However, you are unsure which of the answers A, B, and C is correct. Still, you believe that answer A is more likely to be correct. Therefore, you give answer A two points and answer B and answer C one point.

Option 5
Answer A – 1 point
Answer B – 1 point
Answer C – 1 point
Answer D – 1 point

You are uncertain which answer is correct. You do not feel like one answer is more likely to be correct than the others. Therefore, all questions are given one point.

Rules of the exam:
- Please distribute the four points to the answer(s) you believe suits the question most. If you accidentally distribute more or less than four points, you will not be able to continue to the next question.
- You are not allowed to use calculators and note sheets, or look up answers on the internet.
- During the quiz, you are not able to go back to previous questions.

- This quiz has no time limit.

As mentioned, this research is entirely anonymous. Therefore, read and follow the instructions carefully, because no data can and will be traced back to the participant. But, most importantly, have fun!

You will now continue with the practice question.