Student Learning in a Leaderboard Gamified Micro-lecture: An Experimental Study on the Roles of Achievement Motivation and Perceived Leaderboard Difficulty

Petja O.P. Liimatta

Master Thesis: Learning Sciences

Section of Instructional Technology

Faculty of Behavioural, Management and Social Sciences (BMS)

University of Twente

Supervision and Examination Committee

Dr. Judith ter Vrugte

Dr. Tessa Eijsink

Enschede, April 2023

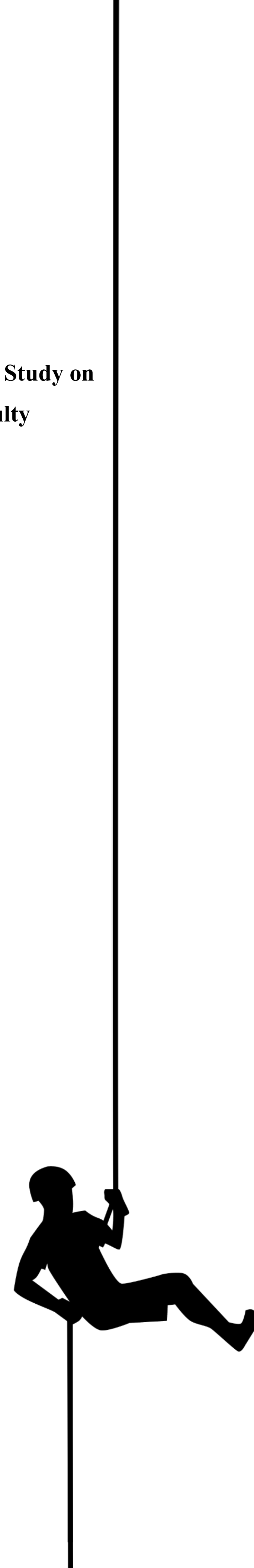


Table of Contents

Table of Contents ... i

Acknowledgements ... iii

Abstract ... v

Student Learning in a Leaderboard Gamified Micro-lecture: An Experimental Study on the Roles of Achievement Motivation and Perceived Leaderboard Difficulty ... 1

Theoretical Framework ... 2

The Integrated Model of Leaderboards Effect on Learning ... 2

Attentional Video Engagement and Learning ... 4

Perceived Leaderboard Difficulty ... 5

Achievement Motivation ... 7

Current Study ... 8

Method ... 10

Design ... 10

Participants ... 10

Materials ... 11

Achievement Motivation (AMS-R) ... 11

Emotion (PANAS) ... 12

Perceived Leaderboard Difficulty ... 12

Attentional Video Engagement ... 13

Description of Micro-lecture ... 14

Domain Knowledge Tests ... 15

The Leaderboard Quiz Platform ... 16

Procedure ... 17

Data Analysis ... 19

Results ... 20

Preliminary Analyses ... 20

Model Test ... 20

Predictors of Learning Performance ... 23

Predictors of Positive and Negative Emotions ... 25

Attentional Video Engagement ... 26

Discussion ... 27

Learning from a Leaderboard Gamified Micro-lecture ... 27

Learning ... 27

Emotions ... 28

Attentional Video Engagement ... 29

The Effects of Perceived Leaderboard Difficulty and Achievement Motivation ... 30

Limitations 32
Implications and Conclusion 33
References 35
Appendix A 41
Appendix B 42
Appendix C 43
Appendix D 46
Appendix E 50
Appendix F 53

Acknowledgements

At the point of writing this, it is extremely difficult to fathom that I am on the edge of transitioning from being a student to becoming a working adult. The closer this edge has come, the more changes have come to my life, with the only stable thing being this thesis. Thus, now writing these Acknowledgements, thanking the people who have allowed me to get this far and nearing the end of my thesis, I must admit I do feel some sort of way.

I must first acknowledge my main supervisor Dr. Judith ter Vrugte for her continuous involvement and adaptability to my ways of working. Her feedback and guidance were not only extremely helpful to shape this thesis but also to support me in the sometimes chaotic thesis writing process. Second, I thank my second supervisor, Dr. Tessa Eijsink for her critical but helpful final remarks now also on my master's thesis. Third, I acknowledge my parents, Linnin and Pasi for their continuous social and financial support. Without them I would not have been able to experience the four and a half years I lived in the Netherlands as I did. Relatedly, I must thank the University of Twente in general for allowing me to start studying psychology in the Netherlands and including the fascinating specialization direction of educational psychology. I originally started at the UT to become a therapist, but thanks to the fascinating courses on learning psychology, I found a direction that fit my interests even more. I am eager to see where this direction takes me next, but perhaps first a short break would be nice.

I must also acknowledge the major help from Jolien van Straalen-Pas for assisting me in the development of the domain knowledge tests used in the experiment of my thesis and allowing my experiment to be included within the research methods course of the first-year psychology bachelor students. Similarly, I thank both Jolien and Amelie Schleich for their major help in allowing me to advertise my study within their research methods courses and lectures. Amelie also helped me with my sometimes-challenging data analysis. I also thank Silas for his immense help with the construction of the draft domain knowledge tests. Additional acknowledgements belong to the BMS Lab for their help and support during the data collection. I also want to thank my mother as well as Julius and Ronja for their feedback throughout the writing process. Especially, I also want to thank Ronja for her assistance with the recruitment of participants but also for her support and company during the data collection. Big thanks also belong to the pilot and main study participants. Other acknowledgements belong to Dr. Hans van der Meij and Karlygash for their introductory assistance with the graasp.eu platform for measuring video engagement. Along with all these acknowledgements, I want to thank both Dr. Anne M. Dijkstra for the fascinating research honours programme, as well as Dr. Andreas Weber for the fascinating conversation I had with him as a part of an assignment within the honours programme. Namely, the main insight I got from this conversation was that “good research is not done alone,” which pursued me to reach out to the aforementioned and other people for help with this personally major work.

I also acknowledge my friends in Finland Masir, Arttu, Ville, Kasperi, and Matias. Even though we have chatted almost every day for these past four and a half years, I have missed you all and it is great to be back Home. Lastly, I acknowledge and thank Seth and Ronja, and my numerous other friends living currently in the Netherlands and Germany. I miss you all and I hope to see you again soon, so we can go bouldering together.

Abstract

Within educational gamification literature, there are calls for further systematic trait-based research on gamification elements. Leaderboard gamification is a prominent gamification element within the literature, and it is a promising method for enhancing student engagement and learning. Following the integrated model of leaderboards' effect on learning, the leaderboard has its effect on learning through mediators, such as emotions and motivation, which in turn are influenced by moderators, such as the individual characteristics of learners and the design characteristics of the leaderboard. Given this framework, the present study developed an experiment to investigate the effects of a leaderboard gamified micro-lecture on students' *emotions, attentional video engagement*, and *learning performance*. The roles of *perceived leaderboard difficulty* and students' *achievement motivation* were also explored. The present study applied a pre–test–intervention–post–test experimental survey design. A final sample of 127 university students were recruited (99 females [78%]; $M_{\text{age}} = 20.46$, $SD = 2.13$). The results from the moderated-mediation models (PROCESS macro), confirmed that a leaderboard gamified micro-lecture is more beneficial for learning performance (domain knowledge post-test performance) than a non-gamified micro-lecture ($b = 0.91$, $p = 0.01$). Against expectations, only negative emotions seemed to be predicted by condition ($b = -1.87$, $p = 0.05$); while in turn, negative emotions predicted learning performance, but only when moderated by high achievement motivation ($b = 0.08$, $p = 0.05$). The effect of perceived leaderboard difficulty was also confirmed, but only when moderated by high achievement motivation ($b = 1.51$, $p = 0.05$). In conclusion, leaderboards can be applied to gamify a micro-lecture to enhance the learning performance of all students, while higher difficulty leaderboards can benefit the learning of a part of the student population.

Keywords: achievement motivation, difficulty, educational videos, emotion, leaderboard

Student Learning in a Leaderboard Gamified Micro-lecture: An Experimental Study on the Roles of Achievement Motivation and Perceived Leaderboard Difficulty

Gamification, “the use of game design elements in non-game contexts” (Deterding et al., 2011, p. 9), can be an effective method to increase students’ motivation, engagement, and academic performance (Manzano-León et al., 2021). According to a recent systematic literature review, many students find the leaderboard as the most engaging gamification element (Zainuddin et al., 2020a). Within the educational context, a leaderboard acts as a ranked visual presentation of the performance of a set of students, which allows performance comparison between students (Christy & Fox, 2014). Many educational studies have found leaderboard gamified learning activities to lead to better learning outcomes (versus non-gamified equivalents) (Ortiz-Rojas et al., 2019; Zainuddin, 2018). However, other research has either found no effects of leaderboards on learning or has found varying motivational, emotional, and engagement outcomes (Frost et al., 2015; Johnson et al., 2020; Zainuddin et al., 2020b). Gamification researchers suggest that these varying findings may result from the interaction between the individual characteristics of the students and gamification design characteristics (Cao et al., 2022; Höllig et al., 2020; Landers et al., 2018). Consequently, the present study joins the calls for such systematic trait-based gamification research (Cao et al., 2022; Höllig et al., 2020; Landers et al., 2018), with the aims of investigating the leaderboard and its effects on student engagement and learning.

As a theoretical framework for the investigation, the present study applied the integrated model of leaderboards effect on learning of Cao et al. (2022). The model argues that a leaderboard prompts differing emotions, motivation, and cognition in learners, which in turn influence their learning performance (mediation effects). Ultimately, the model claims that the leaderboards and mediation effects on learning depend on, for example, the design characteristics of the leaderboard and the individual characteristics of learners (moderation effects) (Cao et al., 2022). Although the model did not involve engagement, Cao and colleagues inferred following their findings that the attention to the learning activity may be an additional relevant factor (Cao et al., 2022). Therefore, the present study sought to investigate this by exploring the role of *attentional engagement* as an additional mediator.

As the learning activity for the investigation, the present study chose to gamify a micro-lecture (i.e. short video lecture; Shatte & Teague, 2020) through a leaderboard. Research suggests that one of the main issues with micro-lectures is with engagement; specifically, micro-lectures are inadequate at encouraging all students to engage with the lecture content (Szpunar et al., 2013; Yang et al., 2021). Since students find leaderboard gamification to be highly engaging (Zainuddin et al., 2020a), leaderboard gamifying a micro-lecture was seen as a possible method to enhance student engagement and learning from such lectures. Thus, combined with the model of Cao et al. (2022), the general research question of the present study was: *to what extent does leaderboard gamifying a micro-lecture influence students' positive emotions, negative emotions, attentional video (or micro-lecture) engagement, and learning when compared to a non-gamified micro-lecture*? Additionally, aligning with the model and trait-based gamification literature (Höllig et al., 2020; Landers et al., 2018), the present study explored the extent to which the leaderboards effects would be influenced by a design characteristic of the leaderboard (*perceived leaderboard difficulty*) and individual characteristic of the students (*achievement motivation*).

Theoretical Framework

The Integrated Model of Leaderboards Effect on Learning

A leaderboard is a visual display of the performance of individuals on a certain activity (Christy & Fox, 2014). Following Cao and colleagues' (2022) integrated model of leaderboards effect on learning, a leaderboard has its main effect on learning through the game mechanics of the leaderboard (e.g., feedback, goal setting, and competition). Namely, the visual display of scores on the leaderboard prompts social comparison between the learners (Bai et al., 2021; Cao et al., 2022; Christy & Fox, 2014; Nebel et al., 2017). These scores provide feedback on one's performance in relation to others, which can encourage learners to set goals for later performance on the leaderboard gamified activity (for instance, to perform better than oneself, others, or both) (Höllig et al., 2020). As such, when compared to a non-gamified counterpart, a leaderboard gamified learning activity prompts individuals to compete for high scores, and this competition has been found to lead to increased learning performance (Plass et al., 2013).

Cao and colleagues' (2022) model argues that the leaderboard's main effect on learning performance is mediated by the *emotions*, *learning motivation*, and *cognition* of learners. In other words, the leaderboard influences the emotions, motivation, and cognition of learners, which in turn impacts their learning performance (Cao et al., 2022). This aligns with previous research, which shows that some students find the leaderboard to be an enjoyable and motivating experience

(Meng et al., 2021; Zainuddin et al., 2020b). Leaderboard research has also found leaderboard gamified activities to result in increased feelings of competence towards the activity (Meng et al., 2021), increased engagement in the learning activity (Zainuddin et al., 2020b), and increased learning performance (Ortiz-Rojas et al., 2019; Zainuddin et al., 2020b). However, at the same time, other students perceive a leaderboard gamified learning activity as a negative experience instead: for example, as a shameful, embarrassing, and discouraging learning experience (Bai et al., 2021; Frost et al., 2015; Tanaka et al., 2016).

The integrated model of leaderboards effect on learning explains such highly varying results through the moderating hypothesis: leaderboards effect is moderated by, for example, *the individual characteristics of the learners* and *the design characteristics of the leaderboard* (Cao et al., 2022). To justify their model, Cao et al. (2022) applied Pekrun's (2006) control-value theory [CVT] of *achievement emotion*. Put simply, the theory claims that individuals experience different emotions and learning motivation in contexts allowing achievement. These emotions and motivation differ based on the interaction between the learning activity and the individual; for example, the emotions are based on the individual's subjective control of the activity and their value of the activity (Pekrun, 2006). Generally, the CVT argues that these prompted emotions can differ by valence (positive or negative). For example, if an individual values an activity highly and expects to perform well in it, then they will experience more positive emotions (such as enjoyment) (Pekrun, 2006), more motivation towards the learning task, and self-report generally higher academic effort (Pekrun et al., 2002). In turn, if an individual does not value a task, or values it but does not expect to do well in it, then they will experience more negative emotions (such as boredom or frustration; Pekrun, 2006), feel less motivated to learn, and self-report generally lower academic effort (Pekrun et al., 2002).

To investigate their model, Cao and colleagues (2022) devised an experiment on the effects of different difficulty leaderboards (a design characteristic of the leaderboard) on learning. Aligning with their model, they explored how the learners reacted to the design characteristic by investigating the learners' perception of leaderboard difficulty (i.e., how easy or difficult do learners find it to be to reach a high rank; Nebel et al., 2017). Additionally, Cao and colleagues (2022) explored the effects of an individual characteristic of the learners, dominant goal orientation (i.e., the preferred achievement goal an individual has in achievement situations, for example, mastery; van Yperen, 2006). Amongst the findings, they discovered that the effects of the leaderboard on learning did not only act through negative emotions and learning motivation (mediation effects) but they were also influenced by the students' perception of leaderboard difficulty (moderation effect) (Cao et al., 2022). Although they did not find positive emotions nor the individual characteristic (dominant goal orientation) to be significant factors (Cao et al., 2022), other research on leaderboards has found individual characteristics of the learners to be significant moderating variables (for instance, achievement motivation; Tanaka et al., 2016). Given their findings, Cao and colleagues concluded that the relations between the leaderboard, emotions, and learning performance are more complicated. As a result, they speculated that the leaderboards may influence learning not only through emotions, but also by the subsequent "concentration of attention" (Cao et al., 2022, p. 12), or attentional engagement to the learning activity, which was to be investigated in the present study.

Attentional Video Engagement and Learning

As the chosen learning activity in the present study was to learn from a micro-lecture, the role of attentional video [or micro-lecture] engagement was explored. *Attentional video engagement* is a sub-dimension of video engagement (Visser et al., 2016). Attentional video engagement is defined as, "attentional focus on the video, with reduced attention to the real world" (Visser et al., 2016, p. 229). Within generative learning theories, it is argued that while the prior knowledge of the learners is the baseline for learning, attention to the learning activity is the first step to learning (Fiorella & Mayer, 2015). Indeed, generative learning theories hold that learning is an active cognitive process (Fiorella & Mayer, 2015, 2016), where the more active the learning process is, the more learning occurs (Chi, 2009). For example, while simply watching an educational video is one of the least beneficial learning methods (passive learning), taking verbatim notes (active learning), making your own notes (constructive learning), and discussing the created notes with others (interactive learning) are more active and thus more beneficial for learning (Chi, 2009). This is because the more the learning material is actively attended, selected, integrated, and organized by the learner, the more learning occurs, because the material becomes part of their long-term memory (Fiorella & Mayer, 2015).

Generative learning theories seem to align with research on achievement emotions. Namely, an achievement activity (such as one gamified through a leaderboard; Cao et al., 2022) may evoke different emotions (positive and negative) and motivation (CVT; Pekrun, 2006), which in turn, according to research on achievement emotions, may influence the attention of the learner on the learning activity (Pekrun et al., 2002). Generally speaking, positive emotions, such as enjoyment and interest, and the accompanying higher motivation, direct the attention of the learner more towards the activity (Pekrun et al., 2002). In turn, low motivation and negative emotions, such as boredom and shame, have a more disengaging effect on attention (Pekrun et al., 2002).

Therefore, as Cao and colleagues (2022) utilized a similar theoretical background for their integrated model of leaderboards effect on learning (for example, Pekrun's CVT), it could therefore be possible that attentional (video) engagement could also be a significant variable (mediating emotions' effect on learning). To further this investigation, it should be explored whether attentional video engagement would also be influenced by the individual characteristics of the learners and design characteristics of the leaderboard, as is argued by the model with regard to the mediation effects of emotions, motivation, and cognition.

Perceived Leaderboard Difficulty

Perceived leaderboard difficulty considers a learners' perspective of leaderboard difficulty (Nebel et al., 2017). *Leaderboard difficulty* is a design characteristic of the leaderboard and it is defined as "how hard or easy [it is] for an individual to get a good rank" or a high rank on a leaderboard (Cao et al., 2022, p. 3). Therefore, with regards to *perceived leaderboard difficulty*, for example, a leaderboard is perceived as low in difficulty when getting a high rank is evaluated to be easy (for instance, due to low performance opponents) and vice versa (Nebel et al., 2017). Generally, gamification research has found that in comparison to a high perceived difficulty leaderboard, learners who perceive a leaderboard to be low in difficulty, experience higher learning motivation (Cao et al., 2022; Nebel et al., 2017), more positive emotions (Cao et al., 2022), and have higher learning performance (Cao et al., 2022; Nebel et al., 2017).

From the perspective of the CVT and achievement emotion literature, the difficulty (or demands) of a learning activity affects emotions, motivation, attention on task, and learning performance by influencing the learner's control perceptions and value of the learning activity (Pekrun, 2006; Pekrun et al., 2002). In other words, a learner initially has a certain value towards a learning activity, and accordingly has a certain expectation of how they will perform in it. However, these values and expectations are influenced by the perceived difficulty of a learning activity, which effect depends on the match between the perceived difficulty and learner's capability. For example, if a learner does not value a learning activity to begin with, they will likely experience boredom (negative emotion) towards the task, regardless of the perception of difficulty (Pekrun, 2006), which in turn leads to lower motivation, more surface-level (or passive) learning strategies, has a disengaging effect on attention towards the learning activity, and has been found to be negatively related to self-reported effort (Pekrun et al., 2002). Alternatively, if a learner does value a learning activity, but perceives it to be too difficult in relation to one's own capabilities, they will experience frustration (negative emotion) and lower motivation towards the learning activity (Pekrun, 2006). In turn, if a learner values a learning activity and perceives it to be reachable due to one's capabilities, they will experience enjoyment (positive emotion) (Pekrun, 2006), which may lead to increased attention towards the learning activity, for instance due to enhanced motivation and the use of more effective learning strategies.

The present study applied the challenge-threat framework of competition of To et al. (2020) to further argue for the effect of perceived difficulty of a learning activity on learning in the context of leaderboard gamification. Similarly to the CVT, the challenge-threat framework argues that the starting point to the effect of a competitive activity on the individual's performance in that activity is the extent to which they are motivated by competitive activities (To et al., 2020). As such, if learners do not find such activities as motivating, only low or null effects occur. Similarly to the CVT, the framework argues that the main effect of competitive activities on performance is dependent on the match between individual's capability and difficulty. Namely, a close match between capability and difficulty, leads to the perception of the task being a doable *challenge*, which motivates the individuals to perform better in the activity. Inversely, even though one would find competitive activities as motivating, an uneven match with too high a perceived difficulty leads to a *threat* perception with regards to the feasibility of the activity, which leads to demotivation and lower performance instead (To et al., 2020). As such, following both the CVT and the challenge threat-framework of competition (Pekrun, 2006; To et al., 2020), it seems that the effects of a leaderboard and perception of leaderboard difficulty may be highly determined by the individual characteristics and capabilities of the learners, which partially aligns with the integrated model of leaderboards effect on learning of Cao et al. (2022).

Achievement Motivation

One such individual characteristic may be *achievement motivation*, which is a relatively stable characteristic of an individual to be motivated by achievement-related activities, which in turn influences their behaviour in such activities (Eccles, 1983). The functions of achievement motivation can be explained through the situated expectancy-value theory [SEVT] of achievement motivation (Eccles & Wigfield, 2020). Generally, the theory argues that the motivation and behaviour of an individual in these achievement-related activities is determined by the interaction between the individual's success expectancy and task value (Eccles & Wigfield, 2020), which are very similar to the constructs in the control-value theory of achievement emotion (Berweger et al., 2022). Consequently, by combining the SEVT and the CVT, it can be argued that the emotions, motivation, and behaviour within an achievement-related activity are partly dependent on how much the individual values the activity and how well they expect to succeed in it (i.e., one's level of achievement motivation; Berweger et al., 2022). Indeed, other research on the achievement motive seems to reflect this (Lang & Fries, 2006). Namely, people are often generalized into two groups with regards to achievement motivation: people more motivated towards succeeding [henceforth, *high achievement motivated individuals*] and people more motivated towards avoiding failure [henceforth, *low achievement motivated individuals*] (Lang & Fries, 2006). High achievement motivated individuals have generally higher performance in such achievement-related tasks, are more persistent in them, enjoy them more, experience more flow in them, evaluate their performance more positively, and set higher goals for achievement (Lang & Fries, 2006). Contrariwise, low achievement motivated students have a more general tendency for worrying, for negative self-evaluation, test-anxiety, they experience less flow, and set lower goals for achievement in such activities (Lang & Fries, 2006).

This description of high achievement motivated students closely resembles the description of the individuals who would be most likely to benefit from competitive environments given by To et al. (2020) within their challenge-threat framework of competition. Namely, "existing research suggests that individuals with personality traits associated with the desire to achieve, enjoyment of competition, and a sense that one has the personal resources needed to attain success exhibit increased intrinsic motivation and performance during competition" (To et al., 2020, p. 22). The results of the experimental study of Tanaka et al. (2016) on leaderboard gamification and achievement motivation partly supports this supposition. Tanaka et al. (2016) discovered that students low in achievement motivation had the lowest motivation to prepare for a leaderboard gamified quiz, while the opposite was the case for high achievement motivated students. Tanaka et al. (2016) concluded that it seems that in comparison to others, low achievement students do not benefit from the motivational aspect of leaderboard ranking.

Current Study

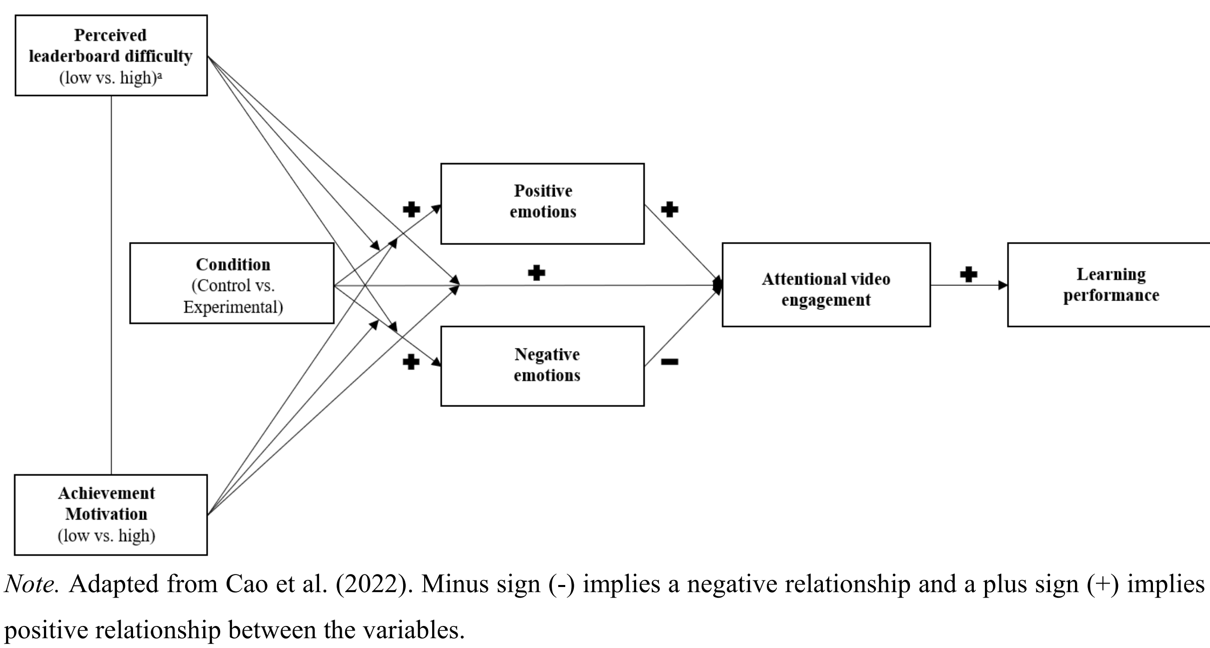
Given the presented theoretical framework, the following research questions were formulated: Firstly, *to what extent does the leaderboard gamified micro-lecture influence learning when compared to the non-gamified micro-lecture?* Secondly, *to what extent are the leaderboard gamified micro-lectures effects on learning mediated by positive emotions, negative emotions, and attentional video engagement when compared to the non-gamified micro-lecture?* On a related note, it was expected that the learning motivation of students would also be influenced, however, due to the limited scope, it was not investigated in the present study. Thirdly, *to what extent do the perceived leaderboard difficulty and students' achievement motivation influence these relations within the leaderboard gamified micro-lecture condition?* With regards to perceived leaderboard difficulty, the present study sought to explore whether simply the perception of leaderboard difficulty would be an influential variable of leaderboards effect. As such, all participants would face a similar leaderboard. In this sense, the present study would differ from previous research (Cao et al., 2022; Nebel et al., 2017), where leaderboard difficulty was clearly manipulated into two different difficulty categories.

To investigate these relations, an experimental design was adopted, where domain knowledge was measured at before (pre-test) and after the micro-lecture (post-test). Participants were randomly divided into either the experimental condition (leaderboard gamified micro-lecture) or the control condition.

With the presented theoretical framework, the following hypotheses were constructed for the experiment (the full proposed moderated-mediation model can be seen in Figure 1). With regards to the first research question on the main effect of the leaderboard, it was hypothesized that *the leaderboard gamified micro-lecture would be more beneficial for learning.* With regards to the second research question on the mediation effects, it was hypothesized that *the effect of the leaderboard on learning would be mediated by positive emotions, negative emotions, and attentional video engagement*. Specifically, it was expected that the leaderboard gamification would lead students to experience more positive, more negative emotions, and yet more engagement. Both the first and the second hypothesis was tested among all participants. Regarding the third research question on the effects of perceived leaderboard difficulty and achievement motivation on these direct and mediation effects, it was first hypothesized that, *within the experimental condition, perceived leaderboard difficulty would influence learning performance*. Namely, due to the demotivating effect of difficulty, those who perceived the leaderboard as high in difficulty (versus low difficulty) would have lower learning performance. As is apparent from the hypothesis, this was tested only among the experimental participants. Secondly, it was expected that *achievement motivation would moderate the effects of leaderboard and perceived leaderboard difficulty on learning performance*. That is, high achievement motivated students in the experimental condition would experience more positive emotions, less negative emotions, and more attentional video engagement, which would lead to higher learning performance in comparison to low achievement motivated students. With regards to perceived leaderboard difficulty, it was expected that low achievement motivated students would be more likely to benefit from the leaderboard when

Figure 1.

The Expected Moderated Mediation Model on the Relations between Condition and Learning Performance



Note. Adapted from Cao et al. (2022). Minus sign (-) implies a negative relationship and a plus sign (+) implies a positive relationship between the variables.

[a] Perceived leaderboard difficulty was only inquired among the experimental condition.

perceived it as low difficulty, while in turn the opposite would the case for high achievement motivated students. As such, this final hypothesis was tested among all participants for leaderboards effect on learning (in comparison to the non-gamified micro-lecture) but also it was tested among the experimental participants for the effect of perceived leaderboard difficulty, which was only inquired among them.

Method

Design

The study followed a pre-test – intervention – post-test experimental design. The experiment involved two conditions, micro-lecture without leaderboard (control) versus a leaderboard gamified micro-lecture (experimental). The on-campus experiment was set up on www.qualtrics.com. In addition to condition, the achievement motivation of participants (low vs. high) and the perception of leaderboard difficulty (low vs, high; only inquired within the experimental condition) were treated as independent variables. During data analysis, condition, achievement motivation, and perceived leaderboard difficulty were used to group the student participants (for example, an experimental condition participant of low achievement motivation and high perceived leaderboard difficulty). In turn, emotions (positive and negative) following the pre-test, attentional video engagement to the micro-lecture, and learning performance (post-test) were treated as dependent variables. Lastly, three variables were inspected as possible control variables: prior knowledge (Cao et al., 2022; Sailer & Homner, 2020), learning interest (Cao et al., 2022), and whether participants took notes during the micro-lecture (a form of active learning; Chi, 2009; Fiorella & Mayer, 2015; Fiorella & Mayer, 2016). The study was approved by a data protection officer (GDPR registration) and the BMS ethical committee of the University of Twente (request nr: 221069).

Participants

University students were recruited in-person through advertisement presentations during research methods lectures and tutorials, online, through the University of Twente [UT] SONA platform, Canvas, social media (WhatsApp, LinkedIn), and via fliers and posters on the UT campus. Registration through the SONA platform rewarded SONA credits upon participation, which bachelor students from psychology and communication sciences needed for graduation. In turn, registration outside of SONA did not give any reward. The inclusion criteria were to be a current university student and to be familiar with research methods.

The sample consisted of 138 participants. Meanwhile, as the present study was interested in comparing participants who clearly rated the leaderboard as either low or high in leaderboard difficulty, the final sample consisted of 127 participants. This is since 11 experimental participants were removed as they perceived the leaderboard as neither low or high in difficulty. The final sample comprised of 99 females (78%) and 28 males, aged 20.46 on average (*SD* = 2.13). Most participants were from Germany (*n* = 64; 50.4%), Netherlands (*n* = 24; 18.9%), and Romania (*n* = 5; 3.9%), while the remainder of the participants were from 23 other countries. All participants completed the whole experiment. Most participants were current UT psychology students (*n* = 94; 74%), of which 71 (55.9%) were current first-year psychology students (2022 cohort).

More specifically, the experimental condition had a total of 81 participants (63.8%), of which 59 were female (72.8%). The experimental participants were aged 20 on average (*SD* = 2). Most of the experimental participants were from Germany (*n* = 40; 49.4%), Netherlands (*n* = 17; 21%), and Romania (*n* = 4; 4.9%). The remainder of the experimental participants were from 14 other countries. Most of the experimental participants were current psychology UT students (*n* = 62; 76.5%), of which 46 (74.2%) were current first-year psychology students (2022 cohort).

Materials

Achievement Motivation (AMS-R)

The revised 10-item version of the achievement motives scale (AMS-R; Lang & Fries, 2006) was employed to measure the *achievement motivation* of the participants (i.e., how individuals are motivated by and behave during achievement-related activities; Eccles, 1983). The AMS-R measures achievement motivation by assessing two dimensions relating to achievement behaviour, *hope of success* and *fear of failure* within an achievement activity (Lang & Fries, 2006). Together, these dimensions measure whether an individual is more afraid of failure within the achievement activity (low achievement motivation) or more hopeful toward success in the achievement activity (high achievement motivation) (Yang et al., 2021). The AMS-R measures the two dimensions with five items each on a 4-point Likert scale (‘4 = strongly agree’; ‘1 = strongly disagree’) (Lang & Fries, 2006). Example items are: ‘I like situations, in which I can find out how capable I am' (hope of success), and 'I am afraid of failing in somewhat difficult situations, when a lot depends on me' (fear of failure) (see Appendix A for the full scale) (Lang & Fries, 2006).

Similarly to Yang et al. (2021) who used the 30-item version of the AMS-R, the final achievement motivation score was computed by subtracting the participant scores on the five high achievement motivation items from the scores on the five low achievement motivation items. A negative score (score between -15 and -1) indicated low achievement motivation, while a positive score (score between 0 and 15) implied high achievement motivation. In previous research utilizing the 30-item version of the AMS-R, internal consistency of .73 has been observed. In the present study, Cronbach's alpha of was .68.

Emotion (PANAS)

The *positive* and *negative emotions* were measured using the positive and negative affect schedule (PANAS; Watson et al., 1988). The PANAS consists of two 10-item scales of positive and negative emotions. It is designed to ask the extent to which participants have experienced specific emotions within a given timeframe. For the present study, the timeframe was adapted to 'right now' (see Appendix B for the adapted PANAS). Accordingly, the participants were asked to rate on a 5-point Likert scale ('1 = very slightly or not at all'; '5 = extremely') to what extent within this specified timeframe they experienced the given positive (e.g., interested, excited, enthusiastic) and negative emotions (e.g., distressed, upset, ashamed).

The total positive and negative emotion scores were calculated by summing all items per scale (minimum score of 10, maximum score of 50). The PANAS has shown good validity and internal consistency regardless of the timeframe, response format, or studied population (Watson et al., 1988). Comparably, strong internal consistencies were found within the present study: namely, alphas were .83 and .80 for the positive and negative emotion scale, respectively.

Perceived Leaderboard Difficulty

Similarly to Cao et al. (2022), the *perceived leaderboard difficulty* was inquired from the participants with one question. Namely, following the leaderboard, the participants were asked to rate on a 9-point Likert scale (1 = very easy; 9 = very difficult),' how difficult did you perceive it to reach a high rank on the leaderboard?'. The score was then used to categorize all experimental participants into three groups. Namely, participants with a score ranging from 1 to 4 were considered to have perceived the leaderboard as low in difficulty, whereas students with a score ranging from 6 to 9 were considered to have perceived the leaderboard as high in difficulty. The remaining students, with a score of 5, were considered to have perceived the leaderboard as neither easy nor difficult (as was designated with the rating on the Likert scale).

Attentional Video Engagement

To measure the *attentional video engagement* (i.e., "attentional focus on the video, with reduced attention to the real world"; Visser et al., 2016, p. 229) to the micro-lecture, it was decided to combine two measures. The decision was based on the engagement literature, which highly suggests that due to the complexity of the construct, multiple measures of engagement should be used in conjunction to both cross-validate and complement each other (Henrie et al., 2015). Consequently, a self-report measure of video engagement was chosen to be used together with an automatic log data gathering platform acting as a video engagement proxy.

Video Engagement Scale (VES). The attention dimension of the video engagement scale of Visser et al. (2016) was selected as the self-report measure. The original 15-item VES focuses on five dimensions of video engagement with three items per dimension: (1) emotions, (2) empathy, (3), identity, (4), attention, and (5) going into a narrative world (Visser et al., 2016). As concluded by Visser et al. (2016), the different dimensions of the original VES can be used separately to measure the different dimensions of video engagement. As the present study was interested in attentional video engagement, only the three items related to attention were applied in the present study. Namely, the participants were asked to rate on a 7-point Likert scale ('1 = 'completely disagree' to 7 = 'completely agree') the extent to which: (1) 'During viewing I was fully concentrated on the video'; (2) 'When I was viewing the video, my thoughts were only with the video'; and (3) 'During viewing, I was hardly aware of the space around me'.

The attentional video engagement score was calculated by summing all three items (minimum score of 3, maximum score of 21). In their original study, the full VES was determined to have good validity and reliability (Visser et al., 2016). In the present study, the Cronbach's alpha for the attentional video engagement dimension was .72, implying good internal consistency.

Unique Play Time. To measure general video engagement during the micro-lecture, the micro-lecture was placed on the www.graasp.eu[1] platform, which, through the VideoPlayer app,

[1] As of the end of 2022 to start of 2023, the graasp.eu platform is to go offline and transferred to graasp.org. To the knowledge of the author, graasp.org will not include the video player app, at least from the start of the launch.

allowed the measurement of the total unique time. Total *unique play time* is the time (in seconds) that a participant watched the micro-lecture at least once. This time does not include replays of the same part of the micro-lecture that was already watched. As described on graasp.eu, unique play time is therefore calculated as “the length of the video minus the sum of the length of the parts that were not played”. As such, “this value can’t be more than the length of the video”.

Similar proxies of video engagement have been used in previous research (Guo et al., 2014; van der Meij & Böckmann, 2021; van der Meij & Dunkel, 2020). However, different from these studies, the present study only measured the unique play time, and not for instance, also the replay time. This decision resulted from the desire to have clear log data. Namely, it was decided that the unique play time data would clearly signify that a participant watched a total of X seconds of the micro-lecture. In turn, if combined with replay data, the acquired total play time could become complicated with regards to interpretation. For example, although a high value within a total play time measure utilizing replay time could signify higher engagement to the micro-lecture, where a participant would rewatch parts of the video they want to understand further. At the same time, the same high value could signify that the participant did not watch the video with proper attention, and thus replayed it. Indeed, the research on such proxies of video engagement utilizing the replay data has not been conclusive. While some previous studies have found replays to be a significant factor (van der Meij & Böckmann, 2021; van der Meij & Dunkel, 2020), others have not (Guo et al., 2014). As such, the present study determined that unique play time would be sufficient to provide a clear picture of the total amount the micro-lecture was watched by a participant, which was thought to act as a proxy of video engagement.

Description of Micro-lecture

For the micro-lecture and the learning material of the present study, research methods, and more specifically the subject of *sampling* was selected. The selected micro-lecture was titled as ‘Sampling’ (https://vimeo.com/135864658). This 12:02 minute long micro-lecture was uploaded to Vimeo by ITC E-learning of the University of Twente. The lecture involves a teacher in front of a PowerPoint presentation. The main aim of the micro-lecture is to act as an introduction to the subject of sampling. The micro-lecture aimed to do this introduction through four learning goals: namely, by introducing (1) when do we need sampling, (2) what does a sampling process look like,

(3) introducing two different types of sampling (probability and non-probability sampling), and (4) what is the sampling bias and the sampling error.

Domain Knowledge Tests

The two domain knowledge tests were based on the selected University of Twente research methods sampling micro-lecture and corresponding Canvas unit for the first-year psychology bachelor students. Based on the material and the adapted guidelines of Gupta et al. (2021) for multiple-choice question [MCQ] test formulation, four learning goals were adapted and used to form two MCQ domain knowledge tests.

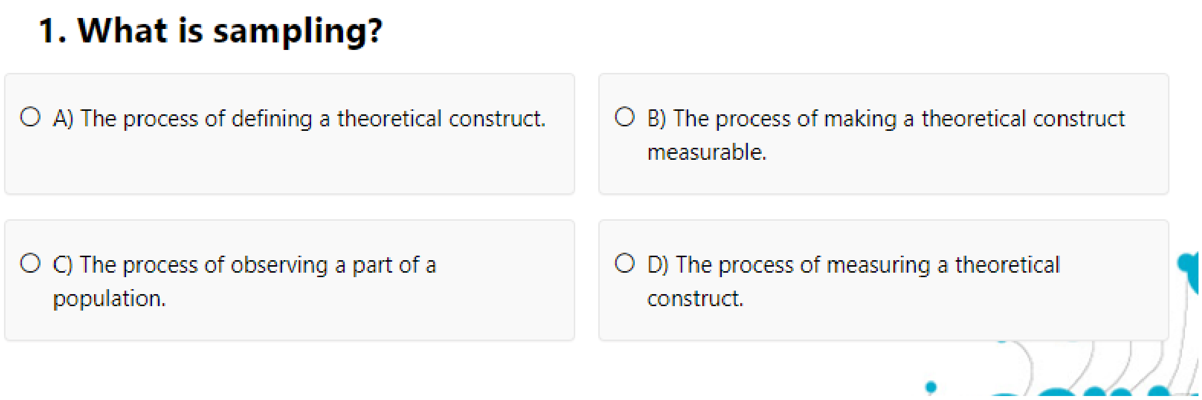
The formulated domain knowledge tests were trialled in a pilot test between the dates of 03/08/2022 and 15/08/2022. The final sample consisted of eight first year psychology students of the university of Twente (cohort 2021), with two being male and six being female. On average, the participants were 21 years of age (*SD* = 1.69). Two of the participants were from the Netherlands, four from Germany, and one from Afghanistan and Poland each. Based on the pilot study, pre-question 3 was modified to be less ambiguous (as no participant answered it correctly), and pre-question 8 was replaced with a new question item (as the previous version was re-assessed not to be sufficiently focused on the subject). Finally, the scores of the eight pilot participants on the pre-test were applied as fake participants to the domain knowledge pre-test leaderboard in the main study as an attempt to standardize participant experience (further explained in *The Leaderboard Quiz Platform*). The learning goals, alongside with the final versions of the domain knowledge pre- and post-tests can be found in Appendix C and Appendix D, respectively

The domain knowledge pre- and post-tests consisted of 10 MCQ items per test. The pre-test questions aimed to measure the prior knowledge of participants, while post-test questions aimed to measure the learning or post-test performance of participants. See Figure 2 for a screenshot of the first MCQ of the pre-test testing the first learning objective (*When do we need sampling?*) on the leaderboard quiz-platform (www.quiz-maker.com).

The scores were calculated by summing all responses to a test (correct answer = 1, incorrect answer = 0), where the higher the score, the higher the prior knowledge or learning (post-test performance) (minimum score of 0, maximum score of 10). With regards to reliability, the Cronbach's alphas were .37 and .35 for the pre- and post-test, respectively, implying poor internal consistency. However, this is to be expected for domain knowledge tests measuring various

Figure 2.

A Screenshot of the First Multiple-choice Question of the Domain Knowledge Pre-test on www.quiz-maker.com



Note. The correct answer is the answer option C.

concepts. Further analyses showed that internal consistency did not improve above .40 upon deletion of any single pre- or post-test question. Therefore, all question items were used in later analyses.

The Leaderboard Quiz Platform

The premium version of www.quiz-maker.com was chosen as the platform for the experiment of the study. The premium subscription was granted for free upon request to use the platform for research purposes. Most essentially for the present study, the platform allows the creation of customizable asynchronous quizzes, the addition of a leaderboard, and automated data gathering.

For both, the control and the experimental condition, the setup was almost identical. Both presented the same 10 MCQs in the same order and with a 10 minute time limit in the bottom of the screen. The time limit was set to control for possible cheating. While the control condition ended with the display of the individual participant's score (score / 10); the experimental version of the quiz showed the leaderboard at the end with a maximum of 30 other participants. The 30 participant setup was due to a limitation with quiz-maker.com, where only the top 30 participant scores would be displayed in a leaderboard. Therefore, six iterations of the experimental quiz were created, which each included eight fake participants as an attempt to set up a standard of experience for all participants as well as to ensure that even the first set of experimental participants would have a comparison group. The fake participants were the scores of the eight pilot participants. At the end of the study the leaderboards included 23 to 25 participants each (including the fake participants). With such a design, the employed leaderboard acted as an absolute leaderboard (which shows all competitors in relation to a participant; Ortiz-Rojas et al., 2019). The participants could observe their score in comparison to that of others on the leaderboard. While rank was not explicitly shown, participants could inquire their ranking by seeing how far they were from the highest scoring participants at the top of the leaderboard.

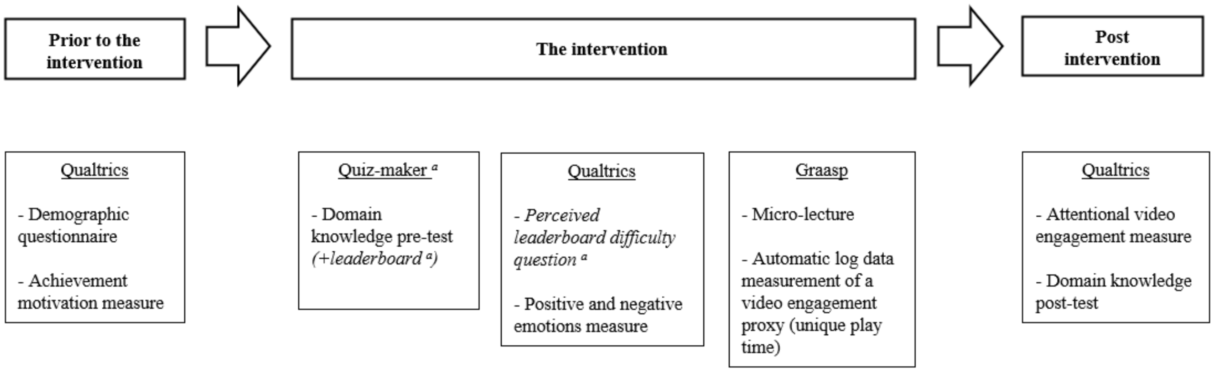
Procedure

The main study was conducted between the dates of 22/09/2022 and 26/10/2022. Participation took approximately 45 minutes. The on-campus experiment involved four to five single-person rooms for participants and one observation room for the author. Each participant room had a computer, headphones, six A4s, a pen, and a folder. The computer had the full experiment open on Qualtrics. The paper, pen, and folder were for participants who wanted to make notes during the micro-lecture. The observation room was set up for the author to observe participants and their screens in case of questions or technical issues, but also to control cheating in the domain knowledge tests.

Each participant was designated a unique ID code to identify them between the platforms. Participants were instructed to ask for help in case of issues and informed regarding the observation. Upon starting the experiment, Qualtrics randomly designated participants to either the control condition or the experimental condition. See Figure 3 for the flowchart of the experiment. In the beginning, the participants were given the informed consent form, demographic questionnaire, and the achievement motivation scale (AMS-R) on Qualtrics. Next, participants were directed to quiz-maker.com, to their respective control or experimental quizzes to test their prior knowledge on sampling. The control participants were given the general instructions to fill out the test honestly and informed about the 10-minute time limit of the pre-test. The experimental participants were given a short introduction page on the leaderboard prior to these instructions. The introduction page was included because research shows that students unaware with leaderboards may experience lower motivational benefits (Tanaka et al., 2016). Following the pre-test, the control participants were shown only their individual score (out of 10), while the experimental

Figure 3

Flowchart of Materials and Platforms used in the Experiment



Note. The underlined names indicate the platforms within which the measures were performed.

[a] Only within the experimental condition.

participants were shown their individual score in relation to the other participants within their leaderboard.

Following the pre-test, participants were directed back to Qualtrics. Here, the experimental participants were first asked to rate to what extent they perceived it to be difficult to reach a high rank on the leaderboard (perceived leaderboard difficulty), after which they were asked to fill out the positive and negative emotion scale [PANAS]. The control participants were only asked to fill out the PANAS. Following, the participants were directed to watch the micro-lecture on graasp.eu, where they were also instructed to freely use the paper and pen in their rooms to make notes on the micro-lecture. While playing the video, the unique play time of each participant was automatically logged by graasp.eu. After the micro-lecture, participants were directed back to Qualtrics and asked to fill the three items from the video engagement scale [VES]. Additionally, all participants were asked to inform whether they had made notes during the micro-lecture (possible control variable), followed by directions to place the made notes into the folder in the room in case they had made some.

Lastly, the participants were given instructions to the domain knowledge post-test measuring their learning (or post-test performance). The instructions were similar from the pre-test, except for the information regarding the timer. Namely, due to differences between the platforms for the pre-test (quiz-maker.com) and post-test (Qualtrics), the timers were different, instead of the whole-test timer applied in the pre-test, the post-test had a timer per question (45 seconds for the first five that were thought to be easier, and 75 seconds for the remaining five, which was in total 600 seconds or 10 minutes). Lastly, a debrief was given to participants revealing their condition and the purpose of the study, followed by the student email of the author.

Data Analysis

The data analyses were performed on SPSS (version 28.0.1.0). The descriptive analysis revealed that the unique play time data (the proxy of engagement) was unusable because (1) the variance showed that almost all participants had watched the complete video, and (2) in some cases, the scores were higher than the video length of 721 seconds (12:02 minutes), implying that the measure was not performed correctly. Therefore, the data was excluded from the analyses. Next, post-hoc Kolmogorov-Smirnov tests showed that most of the variable data were non-normally distributed ($p < .05$). However, normality was the only violated assumption, which is considered the least important when performing linear tests (Gelman & Hill, 2006). As such, the linear tests were still performed.

As a preliminary analysis, prior knowledge, learning interest, and note taking during the micro-lecture were inspected as possible control variables using $\chi 2$ tests and ANOVAs among the students. To achieve this, all participants were categorized into three groups for the control variable analyses. Namely, control, experimental and low perceived leaderboard difficulty, and experimental and high perceived leaderboard difficulty. As perceived leaderboard difficulty was only inquired among experimental participants, this categorization allowed the inspection of differences between the conditions and within the experimental condition in the same analysis, which was of interest within the hypotheses.

Following, to test the hypotheses and the assumed moderated mediation model, the PROCESS macro for SPSS was applied (Hayes, 2013), similarly to Cao et al. (2022). However, one of the preliminary Levene's tests showed significantly different variances between the students in both condition and achievement motivation ($F[5,111] = 2.51$, $p = 0.03$) indicating the F-test to be untrustworthy. Therefore, PROCESS macro was run accounting for the heteroskedasticity using heteroskedasticity-consistent standard error inference (HCSE) (namely, HC3; Hayes & Cai, 2007). With regards to the run models, firstly, to investigate differences between students in the two conditions, a custom moderated mediation model was performed among all students. Here, the two conditions were treated as predictors of learning (post-test performance), while positive emotions and negative emotions were treated as parallel mediators, and attentional video engagement as a serial mediator (to emotions). Achievement motivation was treated as a moderator of the direct and indirect effects. Secondly, to investigate how perceived leaderboard difficulty influenced learning, the same custom moderated mediation model was run only among the experimental participants, with perceived leaderboard difficulty as predictor of learning.

Results

Preliminary Analyses

The descriptive statistics for all control, dependent and independent variables by condition and more specifically within the experimental condition by perceived leaderboard difficulty can be found in Table 1 and Table 2, respectively. As preliminary analyses, possible covariates were explored using $\chi 2$ tests and one-way ANOVAs for whether the student participants differed significantly on note taking, learning interest, and prior knowledge (domain knowledge pre-test score) by the three groups (control [$n = 46$], experimental and low perceived leaderboard difficulty [$n = 37$], and experimental and high perceived leaderboard difficulty [$n = 44$]. The results did not indicate any significant differences between the students among the three groups in note taking, $\chi 2(2) = 2.36$, $p = 0.31$, or learning interest, $F(2, 124) = 0.95$, $p = 0.39$. In turn, the one-way ANOVA showed a significant difference between the three groups of students in prior knowledge, $F(2, 114) = 9.91$, $p < 0.001$. A post-hoc Tukey's HSD test on the three groups of students confirmed that most of the students in the three groups had significantly different means. Namely, participants in the group control condition had significantly lower mean prior knowledge (or domain knowledge pre-test score) from the participants in the group experimental and low perceived leaderboard difficulty ($p = 0.02$). Similarly, participants in the group experimental and low perceived leaderboard difficulty had significantly higher mean prior knowledge from the participants in the group experimental and high perceived leaderboard difficulty ($p < 0.001$). Therefore, as some of the students significantly differed in their prior knowledge, it was treated as a covariate in the following model test.

Model Test

To investigate the assumed moderated mediation model and direct effects, two custom moderated mediation models were created and run using the PROCESS Macro for SPSS through a

Table 1

Descriptive Analyses of the Covariates, Dependent and Independent Variables by Condition

	Control condition (n = 46)	Experimental condition (n = 81)	Total (N = 127)
	M (SD)	M (SD)	M (SD)
Covariates			
Learning interest (min 1 – max 9	6.39 (1.29)	6.30 (1.39)	6.33 (1.35)
Note taking (min 0 – max 1)	0.83 (0.38)	0.83 (0.38)	0.83 (0.38)
Prior knowledge [a] (min 0 – max 10)	7.61 (1.70)	7.73 (1.67)	7.69 (1.67)
Dependent variables			
Positive emotions (min 10 – max 50)	28.91 (7.05)	28.35 (6.46)	28.55 (6.66)
Negative emotions (min 10 – max 50)	15.46 (5.47)	13.78 (4.49)	14.39 (4.92)
Attentional video engagement (min 3 – max 21)	16.37 (2.79)	15.46 (3.55)	15.79 (3.31)
Learning [b] (min 0 – max 10)	7.48 (1.55)	8.02 (1.43)	7.83 (1.49)
Independent variables			
Achievement motivation (min -15 – max 15)	-1.67 (4.94)	-1.21 (4.32)	-1.38 (4.54)
Perceived leaderboard difficulty [c] (min 1 – max 9)	-	4.81 (2.10)	4.81 (2.10)

Note. Control condition = students learning through a micro-lecture. Experimental condition = students learning through a leaderboard gamified micro-lecture.

[a] Domain knowledge pre-test score.

[b] Domain knowledge post-test score.

[c] Only inquired within the experimental condition.

Table 2

Descriptive Analyses of the Covariates, Dependent and Independent Variables within the Experimental Condition by Perceived Leaderboard Difficulty

	Experimental condition, low perceived leaderboard difficulty (n = 37) M (SD)	Experimental condition, high perceived leaderboard difficulty (n = 44) M (SD)
	Covariates	
Learning interest (min 1 – max 9	6.51 (1.12)	6.11 (1.57)
Note taking (min 0 – max 1)	0.76 (0.44)	0.89 (0.32)
Prior knowledge [a] (min 0 – max 10)	8.58 (1.13)	7.02 (1.73)
	Dependent variables	
Positive emotions (min 10 – max 50)	29.27 (6.25)	27.57 (6.61)
Negative emotions (min 10 – max 50)	12.03 (3.80)	15.25 (4.54)
Attentional video engagement (min 3 – max 21)	15.35 (3.75)	15.55 (3.41)
Learning [b] (min 0 – max 10)	8.03 (1.57)	8.02 (1.43)
	Independent variables	
Achievement motivation (min -15 – max 15)	-1.68 (3.82)	-0.82 (4.70)
Perceived leaderboard difficulty [c] (min 1 – max 9)	2.73 (0.96)	6.57 (0.76)

Note. N = 127. Total experimental participant count, n = 81.

[a] Domain knowledge pre-test score.

[b] Domain knowledge post-test score.

[c] Only inquired within the experimental condition.

bootstrapping method while considering the heteroskedasticity of the data and controlling for prior knowledge (domain knowledge pre-test score). The first model (n = 117) investigated the relationship between condition (leaderboard gamified micro-lecture versus control) on learning (domain knowledge post-test score). While also considering the direct and parallel mediations of positive and negative emotions, the direct and serial mediation of attentional video engagement, and moderation effects of achievement motivation (low versus high). The second model (n = 79) utilized the same custom model replacing condition with perceived leaderboard difficulty [PLD] as the predictor, investigating the relationship between PLD and learning performance among the experimental, leaderboard gamified condition participants (see Figure 4 for the compiled findings of both models).

Predictors of Learning Performance

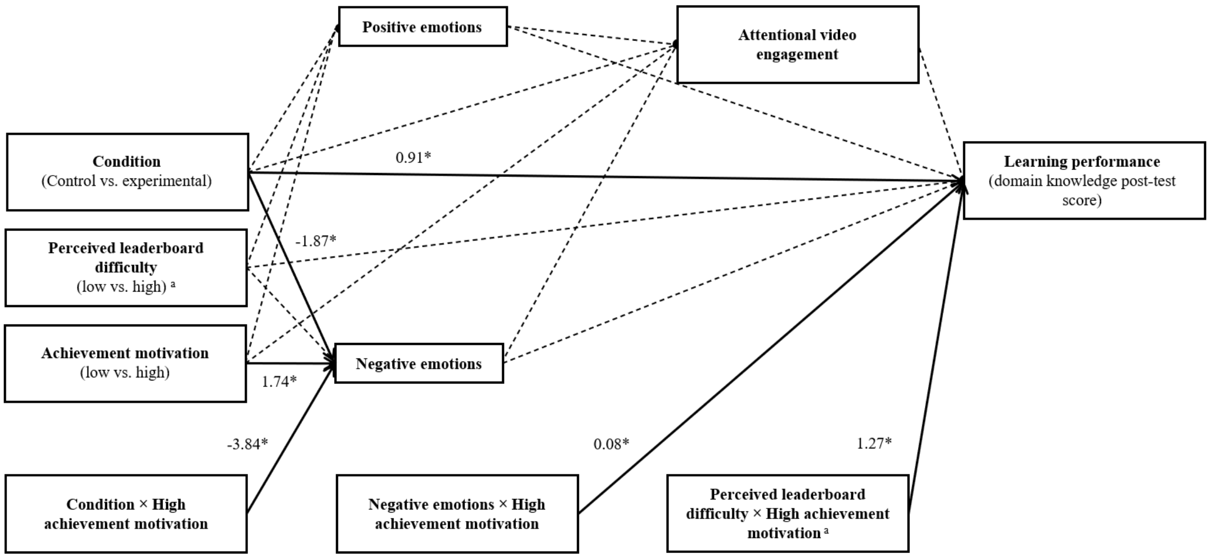
Condition was found to be a predictor of learning performance (domain knowledge post-test score), $b = 0.91$, $F(10, 106) = 2.45$, $p = 0.01$, indicating that the leaderboard gamified micro-lecture led to higher learning performance in comparison to the control condition. Achievement motivation was neither a direct predictor, $F(10, 106) = 2.45$, $p = 0.24$), nor a moderator of the relationship between condition and learning performance, $F(10, 106) = 2.45$, $p = 0.21$.

With regards to the second model, investigating relations between perceived leaderboard difficulty and learning performance among the experimental participants, perceived leaderboard difficulty was not found to predict learning performance, $F(10, 68) = 1.69$, $p = 0.27$. However, when moderated by achievement motivation, perceived leaderboard difficulty was a significant predictor of learning performance, $b = 1.51$, $F(10, 68) = 1.69$, $p = 0.05$. More specifically, higher perceived leaderboard difficulty predicted higher learning performance among high achievement motivated students, $t(68) = 2.24$, $p = 0.03$, with a beta of 1.27, but no effects were observed among low achievement motivated students, t(68) = -0.47, $p = 0.64$ (illustration can be found as Figure E1 in Appendix E).

Returning to the first model, positive emotions did not predict learning performance directly, $F(10, 106) = 2.45$, $p = 0.37$, nor when moderated by achievement motivation, $F(10, 106) = 2.45$, $p = 0.59$. Negative emotions were also not a direct predictor of learning performance $F(10, 106) = 2.45$, $p = 0.73$. However, total negative emotions were found to be a significant predictor of learning performance when moderated by achievement motivation, $b = 0.13$, $F(10, 106) = 2.45$, $p = 0.03$, namely high achievement motivated students had higher learning performance when they experienced more negative emotions, $t(106) = 1.99$, $p = 0.05$, with a beta of 0.08 (can be seen in Figure E2 in Appendix E). However, no effects were observed among low achievement motivated

Figure 4

The Compiled Moderated Mediation Model



Note. $N = 127$. The illustrated model is a compilation of the two performed custom moderated mediation models performed using the PROCESS macro on SPSS. The first model ($n = 117$) investigated the relations between condition (control, non-gamified micro-lecture vs. experimental, leaderboard gamified micro-lecture) on learning performance, while exploring the mediating roles of positive emotions, negative emotions, and attentional video engagement, and the moderating role of achievement motivation. The second custom model ($n = 79$) investigated the experimental condition more closely in terms of perceived leaderboard difficulty on learning performance and the same mediators and moderator. The covariate, prior knowledge (domain knowledge pre-test score) was omitted from the presented model.

[a] Perceived leaderboard difficulty was only inquired among the experimental participants.

$^{*} p < 0.05$.

students, $t(106) = -1.01$, $p = 0.32$. Post-hoc PROCESS moderation models (model 1) on the specific negative emotions showed that the negative emotion, hostility predicted learning performance, revealing that the more hostile participants felt following the pre-test results, the lower their learning performance was in the domain knowledge post-test (the post-hoc moderation models per negative emotions can be found in Appendix F in further detail).

Attentional video engagement was also not found to be a direct predictor of learning performance, $F(10, 106) = 2.45$, $p = 0.24$, nor a predictor when moderated by achievement motivation, $F(10, 106) = 2.45$, $p = 0.81$. All aforementioned findings were supported by both models.

Moving on to inspecting mediation or moderated mediation, the results of the first model suggested the existence of a moderated mediation relationship, where condition predicted learning performance when mediated by negative emotions, but only when moderated by high achievement motivation, *BootB* = -0.32, 95% CIs [-0.882, -0.001], not low achievement motivation, 95% CIs [-0.10, 0.20]. To investigate this further, a post-hoc model 8 using the PROCESS macro was run for the observed moderated mediation effect when controlling for prior knowledge; however, the bootstrapped model did not support the assumption and thus the moderated mediation was rejected. No other mediation or moderated mediation effects were observed in either of the models.

Predictors of Positive and Negative Emotions

Condition was not a significant predictor of positive emotions directly, $F(4, 112) = 3.55$, $p = 0.39$, nor when moderated by achievement motivation, $F(4, 112) = 3.55$, $p = 0.30$. Positive emotions were also not predicted by achievement motivation, $F(4, 112) = 3.55$, $p = 0.44$, nor as shown by the second model, by perceived leaderboard difficulty, $F(4, 74) = 2.13$, $p = 0.74$, or perceived leaderboard difficulty moderated by achievement motivation, $F(4, 74) = 2.13$, $p = 0.47$.

The first model found that negative emotions were significantly predicted by condition, $b = -1.87$, $F(4, 112) = 6.27$, $p = 0.05$. The participants in the experimental, micro-lecture gamified condition experienced more negative emotions than those in the control, non-gamified condition. Post-hoc PROCESS moderation models (model 1) on the specific negative emotions showed that the negative emotion, upsettedness was predicted by condition, $b = 0.27$, $F(4, 111) = 5.91$, $p > 0.01$, indicating that participants felt more upset following the leaderboard showing the domain knowledge pre-test results in the experimental condition. Condition also predicted negative emotions when moderated by achievement motivation, $F(4, 112) = 6.27$, $p = 0.06$. Specifically, it was observed that condition predicted negative emotions when moderated by high achievement motivation, $t(112) = -2.39$, $p = 0.02$, with a beta of -3.84, but not low achievement motivation, $t(112) = -0.18$, $p = 0.86$ (this can be observed in Figure E3 in Appendix E), which was supported by the bootstrapped confidence intervals. Therefore, high achievement motivated students experienced less total negative emotions within the experimental condition and vice versa. In turn, the post-hoc analyses on the specific negative emotions showed that upsettedness and scaredness were predicted by condition when moderated by low achievement motivation, $t(111) = 2.94$, $p > 0.01$, with a beta of 0.27; $t(111) = 1.72$, $p = 0.09$, with a beta of 0.16, which was supported by the bootstrapped confidence. Hence, low achievement motivated students felt more upset and afraid after seeing the leaderboard in the experimental condition, while comparatively, they felt less upset and afraid following their individual pre-test scores in the control condition.

Achievement motivation was also found to be a significant predictor of negative emotions, $b = 1.74$, $F(4, 112) = 6.27$, $p = 0.04$. The higher the achievement motivation of participants, the more negative emotions they experienced following the domain knowledge pre-test scores. More specifically, the post-hoc analyses revealed that achievement motivation predicted the negative emotions upsettedness, $b = 0.39$, $F(4, 111) = 5.91$, $p = 0.03$, scaredness, $b = 0.75$, $F(4, 111) = 4.22$, $p = 0.01$, shame, $b = 0.56$, $F(4, 111) = 3.12$, $p = 0.03$, nervousness, $b = 0.92$, $F(4, 111) = 6.33$, $p = 0.02$, and fear, $b = 0.84$, $F(4, 111) = 3.73$, $p = 0.02$. Therefore, the higher the students were on achievement motivation, the more upset, scared, ashamed, nervous, and afraid they felt during the experiment.

Next, moving back to the second model, perceived leaderboard difficulty was not a direct predictor of negative emotions, $F(4, 74) = 4.21$, $p = 0.14$, nor a predictor of negative emotions when moderated by achievement motivation, $F(4, 74) = 4.21$, $p = 0.14$. The presented findings were supported by the bootstrapped confidence intervals.

Attentional Video Engagement

Condition was not a significant predictor of attentional video engagement alone directly, $F(8, 108) = 1.14$, $p = 0.40$, nor when moderated by achievement motivation, $F(8, 108) = 1.14$, $p = 0.30$. Attentional video engagement was also not predicted by achievement motivation, $F(8, 108) = 1.14$, $p = 0.30$, nor, from the second model, perceived leaderboard difficulty, $F(8, 70) = 1.47$, $p = 0.88$, or perceived leaderboard difficulty moderated by achievement motivation, $F(8, 70) = 1.47$, $p = 0.59$.

Attentional video engagement was also not significantly predicted by positive emotions, $F(8, 108) = 1.14$, $p = 0.09$, positive emotions moderated by achievement motivation, $F(8, 108) = 1.14$, $p = 0.24$, nor negative emotions, $F(8, 108) = 1.14$, $p = 0.72$, or negative emotions moderated by achievement motivation, $F(8, 108) = 1.14$, $p = 0.44$. The findings were supported by the bootstrapped confidence intervals.

Discussion

The present study aimed to advance gamification research through a systematic experiment investigating the effect of leaderboard gamification on micro-lecture *learning performance*. To achieve this, Cao and colleagues' (2022) integrated model of leaderboards effects on learning was applied as the general theoretical framework. In line with the model, the present study utilized the control-value theory to explain the roles of *emotions* (Pekrun, 2006; Pekrun et al., 2002). In addition, the role of *attentional video engagement* was explored through the generative learning theories (Chi, 2009; Fiorella & Mayer, 2015, 2016). The situated expectancy-value theory (Eccles & Wigfield, 2020) and the challenge-threat framework of competition (To et al., 2020) were also applied to explore the role of *achievement motivation* and its interaction with *perceived leaderboard difficulty*.

Aligning with the expectations, the results indicated that students learning through the leaderboard gamified micro-lecture had higher learning performance (domain knowledge post-test score) as opposed to students in the control condition. However, inversely to expectations, students learning through the leaderboard gamified micro-lecture experienced fewer negative emotions versus students in the control condition. As expected, the moderating role of achievement motivation was observed. However, the relation was inverse to expectations, as it was found that the more negative emotions high achievement motivated students experienced, the higher their learning performance was. Furthermore, aligning with expectations, it was observed that high perceived leaderboard difficulty induced higher learning performance and vice versa. However, unexpectedly, perceived leaderboard difficulty was only a significant variable among high achievement motivated students. Regarding the role of achievement motivation between the conditions, it was observed that high achievement motivated students experienced more negative emotions in the non-gamified, control condition, while conversely, they experienced fewer negative emotions in the gamified condition, which aligned with expectations. All findings are discussed below.

Learning from a Leaderboard Gamified Micro-lecture

Learning

The results confirmed that learning through the leaderboard gamified micro-lecture induced higher learning performance (domain knowledge post-test score) than the non-gamified micro-lecture. Although some previous research has been inconclusive on the effects of leaderboard gamification on learning (applied alongside other gamification elements; Frost et al., 2015; Johnson et al., 2020; Zainuddin et al., 2020b), the present finding aligns with other studies where such gamification has been found to stimulate higher learning performance in students (Ortiz-Rojas et al., 2019; Zainuddin, 2018). Among the possible explanations for the varying effects between research and the present study, two possible explanations may be the application of a real leaderboard with real participant scores, (mostly) real student opponents, and controlling for prior knowledge. Aligning with Johnson et al. (2020), simulated leaderboards may induce lower competition among learners, while knowing that participants are competing with peers likely induces more competition, and thus has a more motivational effect on learning. Although the present study did not investigate student perception on the authenticity of the leaderboard score or opponents, this could be a topic for future leaderboard research. Simultaneously, aligning with the present results, prior knowledge is considered to be a highly significant variable that should be controlled for when interested in gamified learning activities (Cao et al., 2022; Sailer & Homner, 2020). The effects of leaderboard gamification on learning via emotions are discussed next.

Emotions

Unexpectedly, only negative emotions were found to be a significant variable, while in turn no mediation effects nor effects of positive emotions were observed. The results implied that the students learning within the leaderboard gamified micro-lecture experienced fewer negative emotions than students in the control condition, but these emotions did not have direct effects on learning performance, which were both contrary to expectations. Although, these null effects of emotions on learning are surprising, negative emotions were found to be a significant predicting variable of student learning when inspecting the moderating role of achievement motivation (discussed in The Effects of Perceived Leaderboard Difficulty and Achievement Motivation). Some of the findings align with the research of Cao and colleagues (2022). Namely, within their research on the integrated model of leaderboards effects on learning, they observed that negative emotions were a significant variable (as well as a mediator) of leaderboards effect on learning, while positive emotions did not have effects on learning (Cao et al., 2022). Interestingly however, the present findings contradict with many other studies within the gamification literature. For example, some studies have observed that learners find leaderboard gamified activities to be generally motivating and enjoyable learning experiences (Meng et al., 2021; Zainuddin et al., 2020b). Yet, other studies have discovered that a part of the student population finds such activities to be a more negative motivational and emotional experience instead (Bai et al., 2021; Frost et al., 2015; Tanaka et al., 2016). As such, the present and previous studies may support the integrated model of leaderboards effects on learning (Cao et al., 2022), where the individual characteristics of learners interact with the leaderboard gamified learning activities' (for example, micro-lectures) effects on learning performance. Namely, the present findings may imply that on its own, a leaderboard gamified micro-lecture might only influence the negative emotions of learners, with no effects of negative emotions on learning on the surface. However, when inspecting the individual differences between the students, only then the negative emotions experienced appear to have an effect on students' learning performance. Additional to the direct and mediating effects of emotions on learning from a leaderboard gamified micro-lecture, the present study also explored those of attentional video engagement, which are discussed next.

Attentional Video Engagement

Attentional video engagement was not found to be a significant variable in leaderboards nor emotions effects on learning, rejecting the hypotheses. Although, these results align with the null results of Johnson et al. (2020), the present findings are surprising as generally leaderboard gamified learning activities have been found to be more engaging than their non-gamified counterparts (Zainuddin, 2018; Zainuddin et al., 2020b). Further, this statement is reinforced by the recent systematic literature review of Zainuddin and colleagues' (2020a), which determined that most students, within the articles of the review, found the leaderboard to be the most engaging gamification element. Encouraged by these promising findings and the guidance provided by previous research (the integrated model of leaderboards effects on learning; Cao et al., 2022; generative learning theories; Fiorella & Mayer, 2015; achievement emotion research; Pekrun et al., 2002), the present study expected that attention to a learning activity, such as attentional video engagement to a micro-lecture, would mediate the effects of leaderboards and emotions on learning. Given the strong foundation for attention and engagement, it is exceptional that no effects of leaderboard and emotions on attentional video engagement, nor attentional video engagement on learning were observed.

An explanation could be that the mere attention or attentional engagement to a micro-lecture may not be a significant factor on leaderboards nor emotions roles on learning. In other words, attention (or attentional video engagement), which is only a dimension of (video) engagement (Visser et al., 2016), may not be a relevant enough factor on its own in terms of indirect effects on, for example, micro-lecture learning. This interpretation could align with generative learning theories, which suggest that attention to learning material is only the first step to learning, which should then be followed by active cognitive processing of the learning material (Fiorella & Mayer, 2015). This interpretation could also align with the integrated model of leaderboards effects on learning, which argues that cognition as a whole mediates leaderboards influence on learning (Cao et al., 2022).

In addition to the discussed mediators, the integrated model of leaderboards effects on learning (Cao et al., 2022) and other research (Höllig et al., 2020; Landers et al., 2018) suggest that the effects of gamification on learning are moderated by the individual characteristics of the learners and design characteristics of the gamification elements. Accordingly, the roles of the individual characteristic, achievement motivation, and the design characteristic of the leaderboard, (perceived) leaderboard difficulty, are discussed next.

The Effects of Perceived Leaderboard Difficulty and Achievement Motivation

Although, perceived leaderboard difficulty was not found to be a significant variable on its own, rejecting the hypotheses; it was however, a significant variable when moderated by high achievement motivation. Aligning with expectations, high achievement motivated students, who perceived the leaderboard as more difficult, had higher learning performance (domain knowledge post-test score) when compared to high achievement motivated students who perceived the leaderboard as lower in difficulty. As such, these findings reinforce previous research that has indicated that perceived leaderboard difficulty is a significant design characteristic of the leaderboards effect on learning (Cao et al., 2022). Similarly, the present findings also support the research that has found achievement motivation to be a significant individual characteristic that moderates leaderboards effect (Tanaka et al., 2016). Consequently, the present study successfully demonstrated that achievement motivation is an individual characteristic that can be added to Cao and colleagues' (2022) integrated model of leaderboards effects on learning, and should be considered in future leaderboard gamification research.

Put together, the results imply that the leaderboard only influenced the learning of high achievement motivated students, which may have acted through their negative emotions. It was observed that high achievement motivated students experienced fewer negative emotions within the leaderboard gamified micro-lecture (versus high achievement motivated students in the control condition), which confirmed the hypothesis. Surprisingly however, the more negative emotions the high achievement motivated students experienced, the higher their learning performance was, which was against expectations. No effects of positive emotions or attentional video engagement were observed between the students on achievement motivation nor perception of leaderboard difficulty, rejecting the hypotheses. Lastly, post-hoc analyses on specific emotions revealed that low achievement motivated students felt more scared when learning from the leaderboard gamified micro-lecture, which partially aligned with expectations; however, unexpectedly the negative emotion did not predict learning performance.

The present findings could be explained by the interaction that occurs between the individual characteristics of the learner and the learning activity (Eccles & Wigfield, 2020; Pekrun, 2006; To et al., 2020). Namely, people differ in how much they value achievement-related activities, how well they expect to achieve in them, and in how much they are motivated by them (situated-expectancy value theory [SEVT]; Eccles & Wigfield, 2020; CVT; Pekrun, 2006; challenge-threat framework; To et al., 2020). In turn, as previously established, the resulting emotions and motivation should influence the learning performance of the learners (Cao et al., 2022). Therefore, the present findings may suggest that high achievement motivated students unsurprisingly value and are more motivated by achievement-related activities such as the leaderboard gamified micro-lecture. In turn, low achievement motivated students may not value nor be motivated by such activities to begin with, which could explain the null effects on learning. This assumption would also align with the conclusion of Tanaka et al. (2016) on their experimental study on achievement motivation and leaderboard gamification. Namely, they suggested that while high achievement motivated students may benefit from leaderboard gamification, low achievement motivated students may only experience adverse or null effects.

Therefore, leaderboard gamification seems to improve the learning motivation (Tanaka et al., 2016), as well as, as shown by the present study, the learning performance of high achievement motivated students. Interestingly however, the leaderboards effect on learning among high achievement motivated students seemed to have acted through negative emotions (though, no mediation effect was observed). This unexpected finding further reinforces the need for future research on emotions to inspect them in further detail than valence (Cao et al., 2022; Pekrun et al., 2002). For instance, emotions also differ by activation level, where some negative emotions may instead have a more motivating and engaging effect on learning performance (Pekrun et al., 2002), which could partially illuminate on the present findings.

Limitations

A possible limitation could have been with the set up of the leaderboards. Specifically, from the start of the data collection, all leaderboards began with fake participants who scored eight (out of 10) on average. This design choice attempting to ensure an equal experience for all experimental participants could have had unintentional effects. Research on leaderboards show that leaderboard difficulty (i.e., how easy or difficult it is to achieve a high rank; Cao et al., 2022) is sometimes determined to be based on opponents' performance or scores (Nebel et al., 2017). As such, the present set up of the leaderboards could have influenced participants' perceptions of leaderboard difficulty, which could have had further unknown effects on the other variables. Future research interested in leaderboard difficulty may benefit from devising distinctly different difficulty leaderboards. For instance, one method could be to set up two leaderboard conditions by manipulating fake participant scores to be four (out of 10) and eight on average, to be more confidently perceived as low and high in difficulty, respectively.

Additional limitations may have been with possibly invasive observation. Specifically, the participants were made aware that they were being observed through a camera in the room to control for cheating in the domain knowledge tests. However, this design may have had unintentional effects, for example on the unique play time data, which causes question on the generalizability of the results. Namely, according to the unique play time data almost all participants watched the whole micro-lecture (with only few seconds of variance), which is unlikely reflective of student learning in a less controlled environment. Thus, future research interested in utilizing such proxies of engagement may benefit from using less invasive observation methods. Alternatively, future research could apply different types of proxies of attention, such as eye tracking, as was suggested by Cao et al. (2022). On a related note, research interested in such real-time, automatic, and less invasive research methods could look into applying methods such as the Empatica E4 wearable, which has been found to be a promising measure of physiological signals. For instance, the Empatica E4 has been used to successfully measure emotions and stress during activities (Cosoli et al., 2021).

Lastly, limitations could have been with regards to the diversity of the sample and participation time variability. Namely, all students did not take part in the experiment at the same time. While some participants had not studied the material before, other participants had either recently or some years back. Relatedly, although most participants were first-year psychology students, many others were from other years of the study, from the master's programme, or even from other studies, such as communication sciences. Therefore, future research should avoid for recruiting too diverse samples.

Implications and Conclusion

One of the main theoretical contributions of the present study was that it demonstrated that all students benefitted from the leaderboards effects on learning. Interestingly, high achievement motivation determined whether perceived leaderboard difficulty was a significant factor. As such, the present study joins the calls for more systematic gamification research investigating specific traits of the learners (Cao et al., 2022; Höllig et al., 2020; Landers et al., 2018) and their interaction per gamification element (Landers et al., 2018). For such research, the present study showed that a useful framework to apply could be Cao and colleagues' (2022) integrated model of leaderboards effects on learning.

In terms of practical implications, the present study illustrated that leaderboard gamifying a micro-lecture can enhance the learning of students. Still, only high achievement motivated students seemed to be influenced by perception of leaderboard difficulty.

Educators or others interested in applying leaderboard gamification into learning activities may feel hesitant as the main means for the leaderboard on learning was shown to be negative emotions. However, the present study signified that, although leaderboard gamification resulted in the experience of more or less negative emotions among many students, these negative emotions either resulted in higher learning performance (among high achievement motivated students) or did not have an influence on learning (among low achievement motivated students). As such, the present study may challenge previous leaderboard gamification research that has not found leaderboards to have an effect on student learning (Frost et al., 2015). Additionally, the present study challenges the conclusions of leaderboard difficulty studies, where generally lower difficulty has been stated to be more beneficial for learning (Cao et al., 2022; Nebel et al., 2017). Indeed, the challenge-threat framework of competition (To et al., 2020) provided a helpful perspective for the present and future research on how different learners might react within competitive (learning) activities. Namely, the present study may imply that developing more difficult (leaderboard gamified) learning activities may be more beneficial for the learning experience of a part of the student population (high achievement motivated students), because they may perceive it as a motivating challenge (To et al., 2020). At the same time, lower difficulty leaderboard gamified learning activities may be more beneficial for the learning experience of other student populations (Cao et al., 2022; Nebel et al., 2017); because for some high difficulty competitive (learning) activities may be too difficult, and as such, be perceived as demotivating threat, lowering their (learning) performance (To et al., 2020). Indeed, following the control-value theory of achievement emotion (Berweger et al., 2022; Pekrun, 2006), the situated expectancy-value theory of achievement motivation (Berweger et al., 2022; Eccles & Wigfield, 2020), and the challenge-threat framework of competition (To et al., 2020), it could be considered that the match between (learner) capability and difficulty of the (learning) activity may be one of the main tenets of such research. However, it must be noted that the present study controlled for prior knowledge, which was a highly significant variable in most analyses. As such, in less controlled or more natural learning environments, different effects of leaderboards may occur. Regardless, based on the present findings, it can be concluded that future researchers, course designers, and educators may feel more at ease to apply innovative leaderboard gamified micro-lectures to effectively enhance student learning within new research and courses.

References

Bai, S., Hew, K. F., Sailer, M., & Jia, C. (2021). From top to bottom: How positions on different types of leaderboard may affect fully online student learning performance, intrinsic motivation, and course engagement. *Computers & Education, 173*, 104297. https://doi.org/10.1016/j.compedu.2021.104297

Berweger, B., Born, S., & Dietrich, J. (2022). Expectancy-value appraisals and achievement emotions in an online learning environment: Within- and between-person relationships. *Learning and Instruction, 77*, 101546. https://doi.org/10.1016/j.learninstruc.2021.101546

Cao, Y., Gong, S.-Y., Wang, Z., Cheng, Y., & Wang, Y.-Q. (2022). More challenging or more achievable? The impacts of difficulty and dominant goal orientation in leaderboards within educational gamification. *Journal of Computer Assisted Learning, 38*(3), 845-860. https://doi.org/10.1111/jcal.12652

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women's math performance. *Computers & Education, 78*, 66-77. https://doi.org/10.1016/j.compedu.2014.05.005

Cosoli, G., Poli, A., Scalise, L., & Spinsante, S. (2021). Measurement of multimodal physiological signals for stimulation detection by wearable devices. *Measurement, 184*, 109966. https://doi.org/10.1016/j.measurement.2021.109966

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

Eccles, J. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75-138). http://education-webfiles.s3-website-us-west-2.amazonaws.com/arp/garp/articles/ecclesparsons83b.pdf

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology, 61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding.* Cambridge University Press. https://tinyurl.com/yx9t6hj3

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*(4), 717-741. https://doi.org/10.1007/s10648-015-9348-9

Frost, R. D., Matta, V., & MacIvor, E. (2015). Assessing the efficacy of incorporating game dynamics in a learning management system. *Journal of Information Systems Education, 26*(1), 59-70. https://aisel.aisnet.org/jise/vol26/iss1/6

Gelman, A., & Hill, J. (2006). Linear regression: The basics. In *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: an empirical study of MOOC videos. In *L@S'14, Proceedings of the first ACM conference on Learning @ scale conference* (pp. 41-50). Atlanta, Georgia, USA. https://doi.org/10.1145/2556325.2566239

Gupta, V., Williams, E. R., & Wadhwa, R. (2021). Multiple-choice tests: A-Z in best writing practices. *Psychiatric Clinics of North America, 44*(2), 249-261. https://doi.org/10.1016/j.psc.2021.03.008

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis : A regression-based approach.* The Guilford Press.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods, 39*, 709-722. https://doi.org/10.3758/BF03192961

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36-53. https://doi.org/10.1016/j.compedu.2015.09.005

Höllig, C. E., Tumasjan, A., & Welpe, I. M. (2020). Individualizing gamified systems: The role of trait competitiveness and leaderboard design. *Journal of Business Research, 106*, 288-303. https://doi.org/10.1016/j.jbusres.2018.10.046

Johnson, C. I., Bailey, S. K. T., & Mercado, A. D. (2020). Does gamification work? Analyzing effects of game features on learning in an adaptive scenario-based trainer. In R.A. Sottilare & J. Schwarz (Eds.), *Adaptive Instructional Systems: Second International Conference, AIS 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22* (pp. 493-504). Springer International Publishing. https://doi.org/10.1007/978-3-030-50788-6_36

Landers, R. N., Auer, E. M., Collmus, A. B., & Armstrong, M. B. (2018). Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming, 49*(3), 315-337. https://doi.org/10.1177/1046878118774385

Lang, J. W. B., & Fries, S. (2006). A revised 10-Item version of the achievement motives scale. *European Journal of Psychological Assessment, 22*(3), 216-224. https://doi.org/10.1027/1015-5759.22.3.216

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*(4), 2247. https://www.mdpi.com/2071-1050/13/4/2247

Meng, N., Lee, S. Y. C., & Chu, S. K. W. (2021). A pilot ethnographic study of gamified english learning among primary four and five students in a rural Chinese primary school. In K. Toeppe, H., Yan, & S. K. W. Chu (Eds.), *Diversity, Divergence, Dialogue: 16th International Conference, iConference 2021*. (pp. 575-586). Springer International Publishing. https://doi.org/10.1007/978-3-030-71292-1_45

Nebel, S., Schneider, S., Beege, M., & Rey, G. D. (2017). Leaderboards within educational videogames: The impact of difficulty, effort and gameplay. *Computers & Education, 113*, 28-41. https://doi.org/10.1016/j.compedu.2017.05.011

Ortiz-Rojas, M., Chiluiza, K., & Valcke, M. (2019). Gamification through leaderboards: An empirical study in engineering education. *Computer Applications in Engineering Education, 27*(4), 777-788. https://doi.org/10.1002/cae.12116

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105. https://doi.org/10.1207/S15326985EP3702_4

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology, 105*(4), 1050-1066. https://doi.org/10.1037/a0032688

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*(1), 77-112. https://doi.org/10.1007/s10648-019-09498-w

Shatte, A. B. R., & Teague, S. (2020). *Microlearning for improved student outcomes in higher education: A scoping review.* OSF Preprints. https://doi.org/10.31219/osf.io/fhu8n

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences, 110*(16), 6313-6317. https://doi.org/10.1073/pnas.1221764110

Tanaka, Y., Uwano, H., Ichinose, T., & Takehara, S. (2016). Effects of gamified quiz to student's motivation and score. In *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)* (pp. 1-4). IEEE. https://doi.org/10.1109/VS-GAMES.2016.7590345

To, C., Kilduff, G. J., & Rosikiewicz, B. L. (2020). When interpersonal competition helps and when it harms: An integration via challenge and threat. *Academy of Management Annals, 14*(2), 908-934. https://doi.org/10.5465/annals.2016.0145

van der Meij, H., & Böckmann, L. (2021). Effects of embedded questions in recorded lectures. *Journal of Computing in Higher Education, 33*(1), 235-254. https://doi.org/10.1007/s12528-020-09263-x

van der Meij, H., & Dunkel, P. (2020). Effects of a review video and practice in video-based statistics training. *Computers & Education, 143*, 103665. https://doi.org/10.1016/j.compedu.2019.103665

van Yperen, N. W. (2006). A novel approach to assessing achievement goals in the context of the 2× 2 framework: Identifying distinct profiles of individuals with different dominant achievement goals. *Personality and Social Psychology Bulletin, 32*(11), 1432–1445. https://doi.org/10.1177/0146167206292093

Visser, L. N. C., Hillen, M. A., Verdam, M. G. E., Bol, N., de Haes, H. C. J. M., & Smets, E. M. A. (2016). Assessing engagement while viewing video vignettes; validation of the video engagement scale (VES). *Patient Education and Counseling, 99*(2), 227-235. https://doi.org/10.1016/j.pec.2015.08.029

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of personality and social psychology, 54*(6), 1063-1070. https://doi.org/10.1037/0022-3514.54.6.1063

Yang, J., Zhang, Y., Pi, Z., & Xie, Y. (2021). Students' achievement motivation moderates the effects of interpolated pre-questions on attention and learning from video lectures. *Learning and Individual Differences, 91*, 102055. https://doi.org/10.1016/j.lindif.2021.102055

Zainuddin, Z. (2018). Students' learning performance and perceived motivation in gamified flipped-class instruction. *Computers and Education, 126*, 75-88. https://doi.org/10.1016/j.compedu.2018.07.003

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020a). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review, 30*, 100326. https://doi.org/10.1016/j.edurev.2020.100326

Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020b). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education, 145*, 103729. https://doi.org/10.1016/j.compedu.2019.103729

Appendix A

The Revised Achievement Motives Scale (AMS-R; Lang & Fries, 2006)

Use the following 4-Point Likert scale (from strongly agree to strongly disagree) to rate the following 10 statements.

1. I like situations, in which I can find out how capable I am. *Measures hope of success.*
2. When I am confronted with a problem, which I can possibly solve, I am enticed to start working on it immediately. *Measures hope of success.*
3. I enjoy situations, in which I can make use of my abilities. *Measures hope of success.*
4. I am appealed by situations allowing me to test my abilities. *Measures hope of success.*
5. I am attracted by tasks, in which I can test my abilities. *Measures hope of success.*
6. I am afraid of failing in somewhat difficult situations, when a lot depends on me. *Measures fear of failure.*
7. I feel uneasy to do something if I am not sure of succeeding. *Measures fear of failure.*
8. Even if nobody would notice my failure, I'm afraid of tasks, which I'm not able to solve. *Measures fear of failure.*
9. Even if nobody is watching, I feel quite anxious in new situations. *Measures fear of failure.*
10. If I do not understand a problem immediately I start feeling anxious. *Measures fear of failure.*

Appendix B

The Brief Measure of Positive and Negative Affect (PANAS; Watson et al., 1988)

This scale consists of a number of words (20 in total) that describe different feelings and emotions. Indicate to what extent do you right now experience each feeling or emotion. Use the following scale to record your answers.

	Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
Interested	○	○	○	○	○
Distressed	○	○	○	○	○
Excited	○	○	○	○	○
Upset	○	○	○	○	○
Strong	○	○	○	○	○
Guilty	○	○	○	○	○
Scared	○	○	○	○	○
Hostile	○	○	○	○	○
Enthusiastic	○	○	○	○	○
Proud	○	○	○	○	○
Irritable	○	○	○	○	○
Alert	○	○	○	○	○
Ashamed	○	○	○	○	○
Inspired	○	○	○	○	○
Nervous	○	○	○	○	○
Determined	○	○	○	○	○
Attentive	○	○	○	○	○
Jittery	○	○	○	○	○
Active	○	○	○	○	○
Afraid	○	○	○	○	○

Appendix C

Domain Knowledge Pre-test Questions on Sampling

LO1 (Learning Objective 1): When do we need sampling?

1. What is sampling?

   A) The process of defining a theoretical construct.
   B) The process of making a theoretical construct measurable.
   C) The process of selecting people from a population to observe.
   D) The process of measuring a theoretical construct.

Answer: C

2. What is the main reason for sampling ?

   A) Any given population is usually too large to be studied as a whole.
   B) To select only the people out of a population that are useful for one's study.
   C) To be able to exclude data that does not fit into one's hypothesis.
   D) Because sampling usually yields better results than using the entire population.

Answer: A

LO2: Distinguish between the two different types of sampling (probability and non-probability sampling).

3. Antonia is interested in pilot testing the exam she created. She asks all of her Dutch friends to trial her test.
   This is an example of ________.

   A) non-probability sampling: because all members of the Dutch population have an equal chance of participating in the study.
   B) non-probability sampling: because NOT all members of the Dutch population have an equal chance of participating in the study.
   C) probability sampling: because NOT all members of the Dutch population have an equal chance of participating in the study.
   D) probability sampling: because all members of the Dutch population have an equal chance of participating in the study.

Answer: B

4. What is an advantage of a random sample?

   A) Random sampling can be very cost effective.
   B) It likely leads to high representativeness of the population.
   C) The time efficiency of drawing a random sample.
   D) Drawing a random sample takes almost no effort.

Answer: B

LO3: Understand the relationships between the population, sampling frame, and sample for probability sampling.

5. Which of the following is most likely NOT an example of a population?

   A) All people in the Netherlands.
   B) German women with multiple sclerosis.
   C) 100 people who have been randomly picked out of a phonebook.
   D) Inhabitants of a single city in Paraguay.

Answer: C

6. What is known as the list from which the sample is drawn from?

   A) Interviewed sample
   B) Population
   C) Units of observation
   D) Sampling frame

Answer: D

7. What is known as the group of individuals observed from a larger population?

   A) Sampling frame
   B) Sample
   C) Interviewed sample
   D) Population

Answer: B

LO4: Understand the concepts of sampling error, sampling bias, and response rate.

8. What mistake is most likely to occur when using a non-probability sampling method such as snowball sampling?

    A) Registration error
    B) Sampling error
    C) Sampling bias
    D) Non-response

Answer C

9. Martijn wants to investigate the mean beard length of Dutch men. He draws a simple random sample of 10 people from the male Dutch population through the population registry. After his study, he concluded that Dutch men have on average 25 cm long beards. However, other Dutch studies found the average is 5 cm.
    Finding an average length that is a lot higher than the average found in other studies is an example of:

    A) Sampling bias: because he only sampled men.
    B) Sampling bias: because he only sampled Dutch people.
    C) Sampling error: because he only sampled 10 people.
    D) Sampling error: because random sampling is not appropriate here.

Answer: C

10. When inspecting her data, Lisa notices that a part of her sample has not answered the questions within her study. She wants to investigate this further by computing the percentage of participants that filled out the questionnaire.
    This is known as computing the ______.

    A) sampling error
    B) response rate
    C) sampling bias
    D) registration error

Answer: B

Appendix D

Domain Knowledge Post-test Questions on Sampling

LO1: When do we need sampling?

1. When conducting research, when do we need sampling?

   A) When we are interested in and able to interview a whole class of students.
   B) When we are interested in and able to interview the experiences of one person.
   C) When we are able to study the whole population that we are interested in.
   D) When we are NOT able to study the whole population that we are interested in.

Answer: D

2. What is/are usually characteristic(s) of a good sample?

   A) A good sample is large in size, but not necessarily representative of the population and randomly generated.
   B) A good sample is representative of the population, but not necessarily large in sample size and randomly generated.
   C) A good sample is randomly generated, but not necessarily large in sample size and representative of the population.
   D) A good sample is representative of the population and large in sample size.

Answer: D

LO2: Distinguish between the two different types of sampling (probability and non-probability sampling).

3. Which of the following is NOT an example of non-probability sampling?

   A) Snowball sampling
   B) Simple random sampling
   C) Convenience sampling
   D) Purposive sampling

Answer: B

4. A researcher asks some close friends whether they would participate in his study. Next to that, the researcher asks them if they know other people that could be asked to participate as well.
   Which sampling method is he using ?

   A) Snowball sampling
   B) Purposive sampling
   C) Convenience sampling
   D) Favour sampling

Answer: A

5. Silas uses a very popular newspaper in the Netherlands to promote participation in his research study.
   This is an example of:

   A) Probability sampling: Silas does not know who reads the newspaper and therefore cannot be biased in the selection of his participants.
   B) Probability-sampling: everyone has an equal chance to purchase and read the newspaper.
   C) Non-probability sampling: everyone has an equal chance to purchase and read the newspaper.
   D) Non-probability sampling: Whether one reads the newspaper or not is not determined by chance but by individual choice.

Answer: D

LO3: Understand the relationships between the population, sampling frame, and sample for probability sampling.

6. Which of the following statements is true regarding the population and the sampling frame?

A) The sampling frame can be larger than the population when the sampling method is used.
B) The population and sampling frame are always the same size.
C) The population is always larger than the sampling frame.
D) While the population contains every individual, the sampling frame only contains those who realistically could be asked to participate.

Answer: D

7. Ronja has conducted an interview study for which she randomly sampled 100 people out of the Dutch population registry. From the sample, she interviewed 67 people and is going through the transcripts of the conversations. However, she has not yet excluded people that have not answered some of her questions.
   Which of the following terms best describes in which stage of sampling the study currently is?

   A) Sampling frame
   B) Sample
   C) Interviewed sample
   D) Data

Answer: C

LO4: Understand the concepts of sampling error, sampling bias, and response rate.

8. A researcher gathered a sample of 500 students for her study. However out of those, only 350 filled out her survey.
   What is the response rate?

   A) 70%
   B) 55%
   C) 30%
   D) 45%

Answer: A

9. Joost wants to investigate how socioeconomic status influences success in the workplace in the Netherlands. He hypothesizes that high social economic status is related to success in the workplace. However, he decides to only sample individuals whose parents were born in the Netherlands because he thinks this will confirm his hypothesis.
   This is an example of:

   A) Sampling error
   B) Sampling bias
   C) Simple random sampling
   D) Discriminatory sampling

Answer: B

10. For sampling in her study, Isabelle uses the population registry of the Netherlands, in which all 17.4 million inhabitants are registered. She randomly selects every 10.000th individual to generate her sample. However, when inspecting her data, she notices that the percentage of Asians in her sample is 5%. As this is a lot less than Isabelle expected, she concludes: “I must have made a mistake; my sample is not representative.”
    What error(s) occurred here?

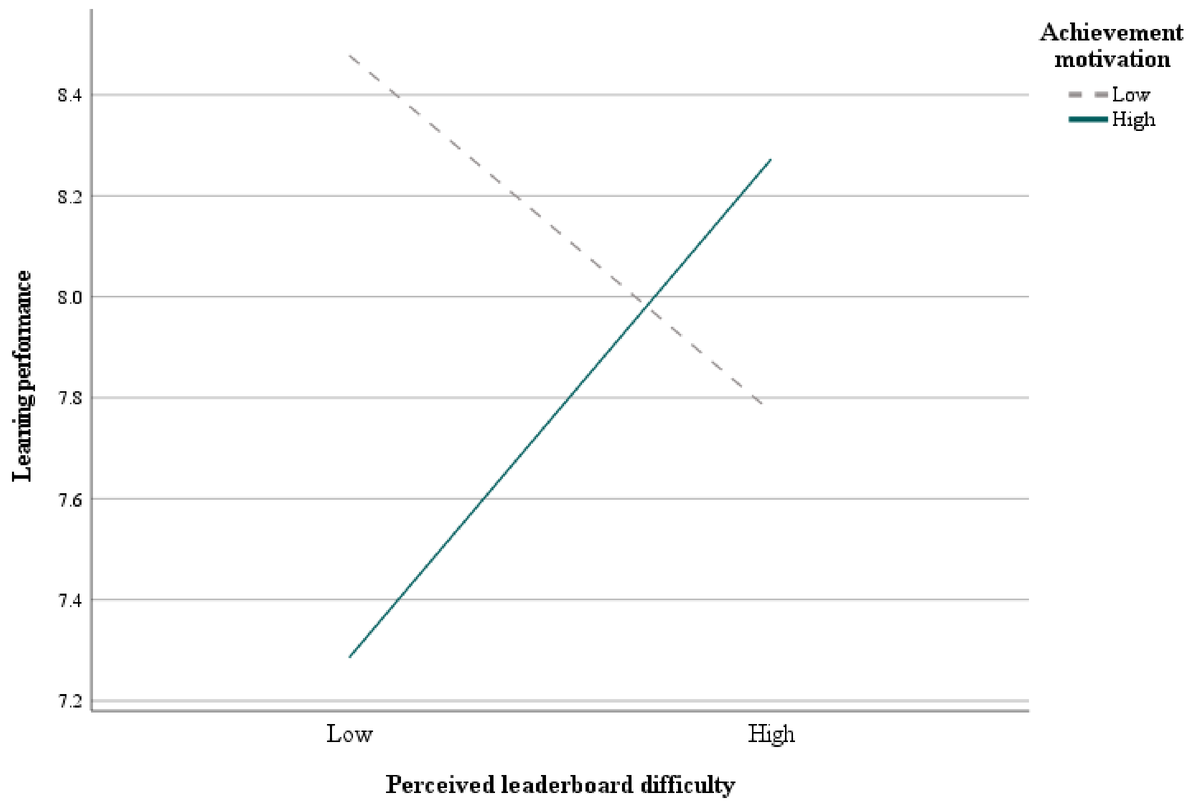
    A) Only sampling bias, but not sampling error.
    B) Only sampling error, but not sampling bias.
    C) Sampling bias and sampling error.
    D) None of the above, the sample should be representative.

Answer: D

Appendix E

Figure E1

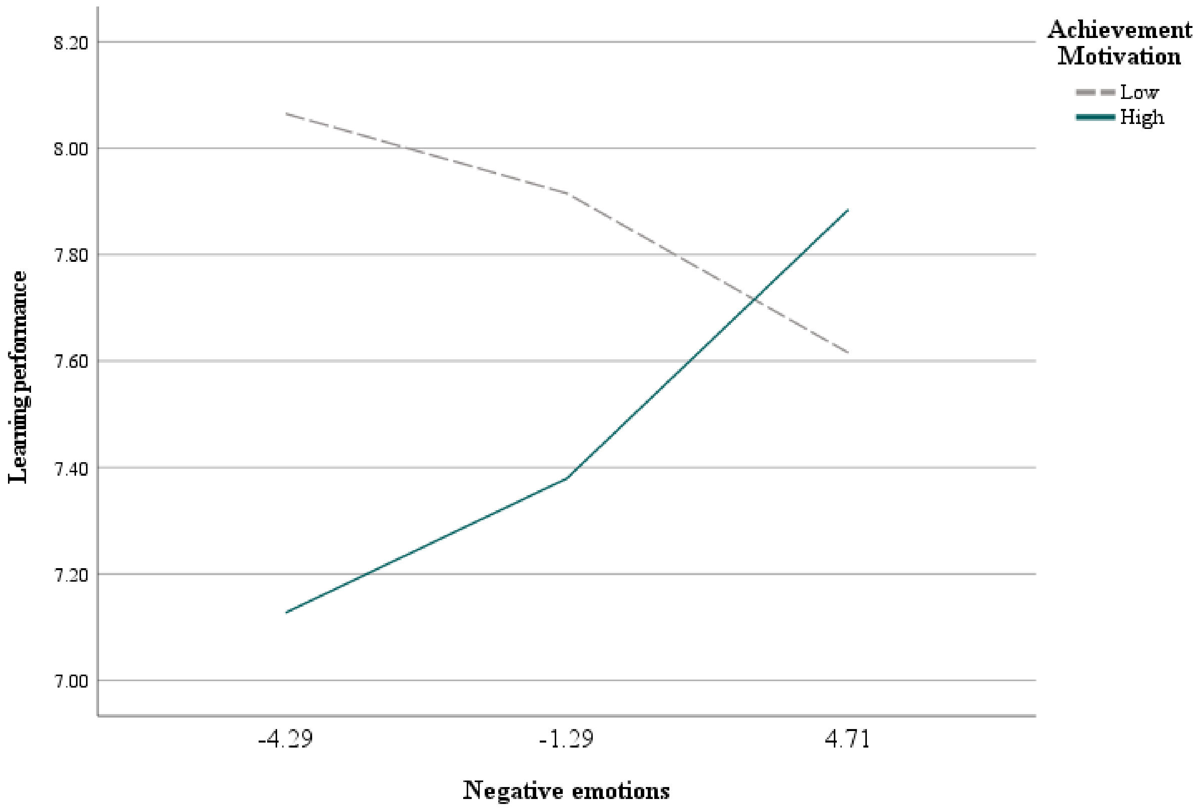
The Effect of Perceived Leaderboard Difficulty on Learning Performance Moderated by Achievement Motivation



Note. n = 79. The figure illustrates the levels of learning performance (domain knowledge post-test score) in the mean centered perceived leaderboard difficulty (low vs high) at -1 and +1 standard deviations from the mean by achievement motivation (low and high). The moderated relationship was only significant ($p < 0.05$) among high achievement motivated students, not those low in achievement motivation.

Figure E2

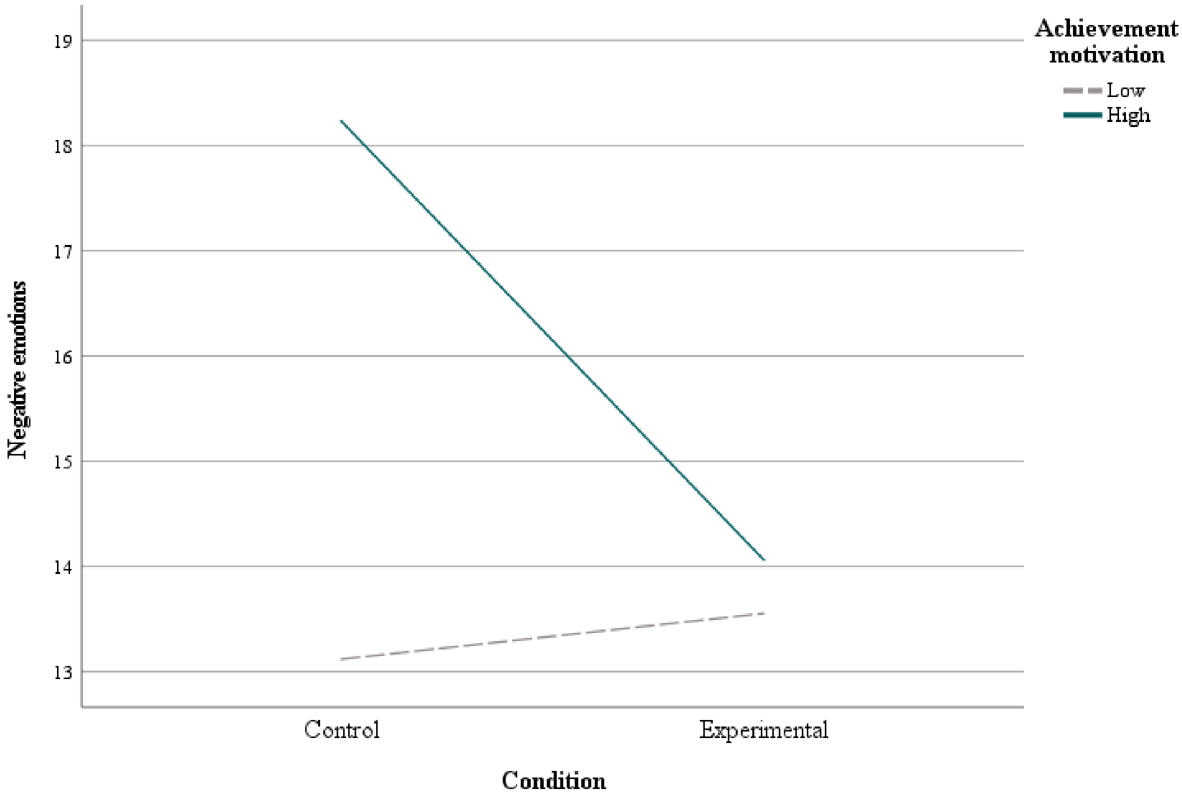
The Effect of Negative Emotions on Learning Performance Moderated by Achievement Motivation



Note. $n = 117$. The figure illustrates the levels of learning performance in te mean centered negative emotions at -1 and +1 standard deviations from the mean by achievement motivation (low and high). The moderated relationship was only significant ($p < 0.05$) among high achievement motivated students, not those low in achievement motivation.

Figure E3

The Effect of Condition on Negative Emotions Moderated by Achievement Motivation



Note. n = 117. The figure illustrates the levels of negative emotions in the mean centered condition (control, non-gamified micro-lecture vs. experimental, leaderboard gamified micro-lecture) at -1 and +1 standard deviations from the mean by achievement motivation (low and high). The moderated relationship was only significant ($p < 0.05$) among high achievement motivated students, not those low in achievement motivation.

Appendix F

To further explore the roles of specific negative emotions, post-hoc analyses were performed per negative emotion (descriptive data can be found in Table F1). 10 separate post-hoc PROCESS Macro moderation models (model 1) were run per negative emotion while controlling for prior knowledge and accounting for heteroskedasticity. Specifically, moderation models were run, where condition predicted each of the negative emotions while moderated by achievement motivation (summary of the findings can be seen in Table F2). Upset was predicted by condition (b = 0.27, p > 0.01), achievement motivation (b = 0.39, p = 0.03), and condition when moderated by achievement motivation (b= -0.49, p = 0.02). More specifically, the moderated relationship occurred only among low achievement motivated students, t (111) = 2.94, p > 0.01, with a beta of 0.27), not those high in achievement motivation, t (111) = -1.13, p > 0.26. Thus, while in general participants felt more upset in the experimental condition, the higher the participants' achievement motivation was, the more upset they also felt after the domain knowledge pre-test. However, with regards to condition, only low achievement motivated students felt more upset in the experimental condition and vice versa.

Scared was also predicted by achievement motivation (b = 0.75, p = 0.01) and condition when moderated by achievement motivation (b = 0.27, p = 0.06), as supported but the bootstrapped confidence intervals. Further, as the bootstrapped confidence intervals did not include zero, the moderated relationship was interpreted to only occur among low achievement motivated students, t(111) = 1.72, p = 0.09, with a beta of 0.16, but not high achievement motivated students, t(111) = -1.44, p = 0.15. Therefore, the higher the students were on achievement motivation, the more scared the felt in general after the domain knowledge pre-test. Though, when inspecting individual differences between conditions, only low achievement motivated students felt more scared in the experimental condition (in comparison to the control condition). Next, achievement motivation was found to predict the negative emotions ashamed (b = 0.56, p = 0.03), nervous (b = 0.92, p = 0.02), and afraid (b = 0.84, p = 0.02). As such, the higher the students were on achievement motivation, the more ashamed, nervous, and afraid they felt after the domain knowledge pre-test.

Lastly, to explore the specific negative emotions that predicted learning performance when moderated by achievement motivation, 10 custom PROCESS macro models were per negative emotion as the predictor of learning performance, while controlling for prior knowledge. No

Table F1

Descriptive Analyses for the Relations Between Condition and Achievement Motivation per Negative Emotion

		Condition		Achievement motivation		High achievement motivation × Condition	
Negative emotion	Total	Control	Experimental	Low	High	Control	Experimental
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Distressed	1.88 (1.06)	1.93 (1.04)	1.85 (1.08)	1.83 (1.18)	1.95 (0.90)	2.10 (0.94)	1.86 (0.88)
Upset	1.25 (0.55)	1.22 (0.51)	1.26 (0.57)	1.19 (0.52)	1.32 (0.58)	1.48 (0.68)	1.23 (0.49
Guilty	1.17 (0.52)	1.17 (0.44)	1.16 (0.56)	1.11 (0.44)	1.23 (0.60)	1.29 (0.56)	1.20 (0.63)
Scared	1.37 (0.72)	1.43 (0.91)	1.33 (0.59)	1.19 (0.46)	1.59 (0.91)	1.90 (1.18)	1.40 (0.65)
Hostile	1.37 (0.73)	1.36 (0.78)	1.32 (0.70)	1.44 (0.81)	1.28 (0.62)	1.43 (0.75)	1.19 (0.52)
Irritable	1.45 (0.74)	1.50 (0.84)	1.43 (0.69)	1.30 (0.55)	1.64 (0.90)	1.76 (1.09)	1.57 (0.78)
Ashamed	1.25 (0.65)	1.37 (0.85)	1.17 (0.50)	1.11 (0.36)	1.41 (0.87)	1.71 (1.15)	1.23 (0.60)
Nervous	1.83 (1.10)	2.11 (1.20)	1.67 (1.02)	1.59 (0.99)	2.14 (1.17)	2.57 (1.21)	1.89 (1.08)
Jittery	1.60 (0.83)	1.84 (0.98)	1.47 (0.71)	1.49 (0.70)	1.75 (0.97)	2.20 (1.15)	1.49 (0.74)
Afraid	1.35 (0.84)	1.46 (1.03)	1.29 (0.72)	1.17 (0.61)	1.57 (1.02)	1.90 (1.37)	1.37 (0.69)

Note. N = 127. Control = non-gamified micro-lecture (n = 46), Experimental = leaderboard gamified micro-lecture (n = 81), Low achievement motivated students (n = 70), high achievement motivated students (n = 57), high achievement motivated students in the control condition (n = 21), high achievement motivated students in the experimental condition (n = 36). The participants scored each individual negative emotion 5-point Likert scale ('1 = very slightly or not at all'; '5 = extremely') on the extent to which they experienced the given emotion 'right now' after the domain knowledge pre-test and shortly prior to the micro-lecture.

moderated relationships were found. However, hostility was found to directly predict predicted learning performance (b = -0.71, p < 0.01). Hence, the more hostile the students felt after the domain knowledge pre-test, the lower their learning performance was in the post-test.

Table F2

Summarized Results of the Moderated Mediation Models with Condition, Achievement Motivation, and Learning Performance per Negative Emotion

	Condition on Negative emotions	Achievement motivation on Negative emotions	Condition × Achievement motivation on Negative emotions	Negative emotions × Achievement motivation on Learning
Negative emotion	*Beta*	*Beta*	*Beta*	*Beta*
Distressed	-0.10	-0.05	0.03	0.18
Upset	**0.27****	**0.39***	**-0.49*** [b]	0.04
Guilty	0.06	0.18	-0.15	0.58
Scared	0.16	**0.75***	**-0.61**[a]	0.09
Hostile	-0.16	-0.16	-0.01	0.45
Irritable	-0.02	0.44	-0.15	0.44
Ashamed	-0.03	**0.56***	-0.50	-0.44
Nervous	-0.26	**0.92***	-0.53	0.04
Jittery	-0.09	0.56	-0.06	0.40
Afraid	0.03	**0.84***	-0.66	0.14

Note. n = 116. The shown results are from 20 independent moderation models run through the PROCESS Macro on SPSS. 10 of the models explored how condition predicted each of the negative emotions while moderated by achievement motivation (results are the first three columns), while in turn, the other 10 moderation models explored how each of the negative emotions predicted learning performance while moderated by achievement motivation (last column).

Significant results are marked in Bold.

[a] Shown to be significant due to bootstrapped confidence intervals not including zero.

[b] Further analyses showed that the relation was only significant among low achievement motivated students.

$^{*}p < 0.05$, $^{**}p < 0.01$.