Bai, S., & Hew, K. F. (2025). How different gamified leaderboards affect individual students' learning engagement, strategies, performance, and perceptions in online classes. *Educational Technology & Society*, *28*(4), 42-61. https://doi.org/10.30191/ETS.202510_28(4).RP03

How different gamified leaderboards affect individual students' learning engagement, strategies, performance, and perceptions in online classes

Shurui Bai[1*] and Khe Foon Hew[2]

[1]Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR // [2]Faculty of Education, The University of Hong Kong, Hong Kong SAR // tstbai@eduhk.hk // kfhew@hku.hk

[*]Corresponding author

(Submitted March 28, 2024; Revised April 21, 2025; Accepted April 26, 2025)

ABSTRACT: Leaderboards visually display social comparison results. The effects of specific situational factors embodied by absolute and relative leaderboards have received little attention. In this study, we compared the effects of an absolute and a relative leaderboard on students' learning engagement, strategies, performance, and perceptions in fully online classes. The results showed that, compared with the absolute leaderboard group, comparison with neighboring competitors in the relative leaderboard group led to a higher level of learning engagement and performance, and encouraged constructive competitiveness that motivated the students to prioritize knowledge mastery over mere task completion. The students at all levels in the relative leaderboard group reported a satisfied, motivated, and progress-oriented attitude. Although the top-ranked students in the absolute leaderboard group were motivated to work harder and were happy to see their achievements publicly displayed, the middle- and bottom-ranked students reported increased peer pressure as they progressed. Our findings lead to the following two practical implications. First, use relative leaderboards to foster constructive competition. Second, limit the public display of the rankings of bottom-ranked students to enhance their learning motivation and engagement.

Keywords: Absolute leaderboard, Relative leaderboard, Gamification, Social comparison

1. Introduction

Many studies have examined the role of gamification in online student engagement (Park & Kim, 2021). Gamification involves using game elements in non-game settings (Deterding et al., 2011). Although gamification aims to promote active learning, research findings on its impact on engagement and learning performance remain mixed (Reyssier et al., 2024). Many scholars have cautioned that gamification does not guarantee positive learning outcomes (Schöbel et al., 2023), as students often respond differently to various game elements (Lavoué et al., 2019), influenced by their unique expectations and preferences (Bai & Liu, 2024; Bennani et al., 2022). To determine when and how a gamification design is effective, it is essential to examine the effects of individual game elements, rather than multiple elements simultaneously, as combining them would confound the results (Gao, 2024).

In this study, we examine the use of leaderboards. A leaderboard is a game element commonly used in gamification to compare an individual's performance, motivation, and engagement with their peers (Li et al., 2024). As a visual representation of social comparison results, leaderboards are often used to elicit a sense of competition to increase student effort (Do et al., 2024). Zichermann and Cunningham (2011) categorized leaderboards into two broad configurations: infinite (or absolute) and non-disincentive (or relative) types. Absolute leaderboards indicate the exact positions of all players and are commonly used in educational contexts (Do et al., 2024). In contrast, relative leaderboards only show a specific user's immediate neighbors in the ranking (Ortiz-Rojas et al., 2025). For example, a relative leaderboard may show up to four neighbors, two above and two below the focal user.

Previous interventions involving leaderboards have largely focused on absolute leaderboards (Li et al., 2024). This approach has left a notable gap in the literature: the lack of direct and controlled comparisons between different leaderboard configurations such as absolute and relative leaderboards. Consequently, educators lack evidence-based insights into which leaderboard configuration can support student engagement and learning more effectively. The importance of addressing this gap is underscored by the recent systematic review conducted by Li et al. (2024) who identified leaderboard configuration as a key determinant of its' effectiveness. Therefore, addressing this gap is crucial for designing leaderboards that enhance engagement and learning outcomes.

The rationale for this comparison lies in the possible distinct motivational effects of different leaderboard configurations on learners. Although absolute leaderboards can foster a sense of competition by revealing the position of all participants, they risk discouraging lower-ranked students. Moreover, players at the top may have a greater sense of achievement than those at the bottom (Ortiz-Rojas et al., 2025). In contrast, relative leaderboards can mitigate feelings of inadequacy by focusing on immediate peers. Specifically, relative leaderboards can alleviate the feeling of demotivation of lower-ranked players (Do et al., 2024). Without understanding these nuances, gamified systems risk undermining engagement rather than strengthening it. For example, in an absolute leaderboard, a struggling student might disengage after seeing their position at the bottom, while a relative format might motivate them by highlighting progress relative to similar peers.

This study aims to fill this research gap by directly comparing the effects of absolute and relative leaderboards on student engagement and learning outcomes. In doing so, we can clarify whether disparities in student engagement and learning performance stem from the inherent characteristics of leaderboard types (e.g., full transparency of absolute leaderboards vs. bounded social comparison in relative leaderboards). The findings can provide actionable insights for educators and gamification designers, enabling them to implement leaderboards in ways that enhance their effectiveness in educational settings.

2. Literature review

2.1. Related studies

To date, the results regarding the effects of leaderboards on students' learning performance remain inconclusive (Li et al., 2024). On the one hand, several studies have suggested that leaderboards can induce better learning performance than a non-gamified condition (e.g., Lee et al., 2022; Ortiz-Rojas et al., 2025; Pi et al., 2024). For instance, Lee et al. (2022) compared the learning performance of a gamified group using a team-based absolute leaderboard and a non-gamified group. Their results indicated that the gamified group performed better than the non-gamified group. On the other hand, other studies have reported no significant effects of leaderboards on learning (e.g., Balci et al., 2022; Cigdem et al., 2024). For example, Cigdem et al. (2024) found a non-significant difference between the leaderboard group and a non-gamified group on learning performance.

Similarly, the potential of leaderboards to promote engagement shows conflicting results. Although some studies have reported that leaderboards enhance engagement compared with non-gamification (Landers & Landers, 2014; Nadeem et al., 2023), others have found non-significant effects on engagement (Cigdem et al., 2024; Ortiz-Rojas et al., 2019; Philpott & Son, 2022). For example, Nadeem et al. (2023) observed that the leaderboard group outperformed non-leaderboard groups in completing course activities. In contrast, Philpott and Son (2022) reported no significant enhancement in student completion of course activities in the presence of a leaderboard.

These conflicting findings highlight the complexity of leaderboard designs and the need for a deeper understanding of the factors that influence their effectiveness. Comparative studies have used a wide range of leaderboard configurations, such as absolute leaderboards, team and individual leaderboards, leaderboards presenting both points and ranking, leaderboards showing only ranking, anonymous leaderboards, artificial leaderboards, and Top N leaderboards, which only display the rankings of the top few users (Li et al., 2024). To date, most studies have used the absolute leaderboard configuration. Although Top N leaderboards may resemble absolute leaderboards, they exclude the majority of users from visibility. This limitation can skew the results by omitting critical data about the broader user base and obscuring the true impact of leaderboards on student engagement and learning. Furthermore, relative leaderboards, which focus on local comparisons and may reduce discouragement among lower-ranked users, remain underexplored. Additionally, many studies have compared leaderboards with non-leaderboard conditions, which may inherently favor leaderboards due to the motivational appeal of achieving a high ranking (Hew & Lee, 2019). Such comparisons limit our understanding of how specific leaderboard designs influence educational outcomes.

To address these limitations, studies that directly compare specific leaderboard configurations, such as absolute and relative leaderboards, are needed to understand how they influence student engagement and achievement. Notably, no published study has directly compared absolute and relative leaderboards (Li et al., 2024) under controlled conditions (e.g., using the same course and instructor) to minimize confounding variables. By going beyond simplistic comparisons with non-gamified conditions and investigating the nuances of leaderboard design, researchers can provide clearer insights into their educational potential.

2.2. Theoretical background

Social comparison theory posits that individuals have an innate desire to evaluate their own abilities (Festinger, 1954), often by comparing their performance with that of others (Yang et al., 2023). In educational settings, social comparison can lead to constructive or destructive competition. Constructive competition fosters a positive environment in which students are motivated to develop their abilities and gain a sense of accomplishment (Sheridan & Williams, 2011). This type of competition enhances learning outcomes by encouraging students to engage deeply with the material. In contrast, destructive competition occurs when the competitive environment leads to increased stress, anxiety, and fear of failure (Houshmand, 2015). This can discourage lower-performing students and create a sense of disengagement, ultimately hindering their learning and well-being.

Leaderboards, a widely used tool to foster competition (Slamet et al., 2024), can influence whether competition becomes constructive or destructive. Two key situational factors—proximity to a performance standard and the number of visible competitors—can play a critical role in shaping competition (Garcia et al., 2013). Absolute leaderboards, which display the scores of all participants, often highlight major gaps between top- and bottom-ranked individuals. While top performers may feel a sense of achievement due to their proximity to the standard (e.g., number one ranking, Ortiz-Rojas et al., 2025), lower-ranked individuals may perceive the gap as insurmountable, leading to feelings of inferiority and demotivation (Qiao et al., 2024). This is largely due to the natural tendency of individuals to engage in upward comparisons, that is, to compare themselves with those who perform better (Gerber et al., 2018). When the gap appears too large, upward comparisons may erode motivation rather than inspire it (Wang et al., 2024), as lower-ranked individuals may believe they have no realistic chance of closing the gap (Na & Han, 2023). As a result, absolute leaderboards may inadvertently exacerbate destructive competition, as those at the bottom may experience increased stress and disengage altogether, believing that improvement is impossible.

In contrast, relative leaderboards, which display only the closest neighbors (i.e., peers with similar performance levels), may promote constructive competition by addressing these situational factors. By reducing the number of visible competitors, they alleviate the feeling of overwhelm (Lessel et al., 2022) and reduce the scope of comparison. This localized comparison enables participants to perceive themselves as close to a performance standard, as they can only compare themselves with peers immediately above and below them, even if they are not near the top of the overall ranking (Bai et al., 2021). This proximity to a standard fosters a sense of achievable progress and reinforces the urgency of self-improvement, as the perceived gap to the next achievable milestone becomes more manageable and less intimidating. By eliminating the daunting prospect of competing against a large and seemingly insurmountable group, relative leaderboards may foster a constructive competitive environment that motivates individuals to engage more deeply and strive for incremental progress.

Given these insights, we propose that relative leaderboards are more likely to foster constructive competition than absolute leaderboards. This difference is expected to influence learners' engagement, strategies, and performance. Specifically, we propose the following hypotheses:

- Hypothesis 1: Learners in all positions of a relative leaderboard exhibit higher levels of learning engagement than those in an absolute leaderboard.
- Hypothesis 2: In a relative leaderboard, learners adopt different learning strategies than in an absolute leaderboard.
- Hypothesis 3: Learners in all positions of a relative leaderboard achieve better learning performance than those in an absolute leaderboard.

Therefore, we addressed the following research questions:

- Research question 1: How do absolute and relative leaderboards affect students' learning engagement in fully online learning?
- Research question 2: How do absolute and relative leaderboards affect students' problem-based learning strategies in fully online learning?
- Research question 3: How do absolute and relative leaderboards affect student performance in fully online learning?
- Research question 4: How do students perceive the use of absolute and relative leaderboards in fully online learning?

3. Method

In this study, we used a quasi-experimental research design combined with qualitative research to investigate the effects of different leaderboard types on students' learning engagement, strategies, performance, and perceptions in fully online classes. The sample included postgraduate students enrolled in an e-learning design course (10 weeks), divided into two groups based on the type of leaderboard used: absolute or relative. The data collection methods included trace data from student interactions, pre-test and post-test assessments, and open-ended surveys. For data analysis, we used epistemic network analysis (ENA) to understand how the students' learning strategies were connected in a problem-based learning environment.

Experiment 1 was conducted in the Fall 2021 semester (Sep–Dec 2021) and Experiment 2 in the Spring 2022 semester (Jan–May 2022). The teaching team, learning materials, course design, assessments, and rubrics were the same regardless of the type of leaderboard used. We used the Moodle learning management system to manage all learning content and gamified activities. The aim and procedures of the study were clearly explained to all participants during the first lesson. Ethical approval for this study was granted to the first author by the university. Only those who signed the consent form were included in this study.

3.1. Setting up the leaderboards

The "Level Up" Moodle plugin was used to set up the absolute and relative leaderboards (Sinnott & Xia, 2020). A review conducted by Sinnott and Xia (2020) evaluated this plugin in terms of its effects on students' learning performance and suggested that it is a suitable application for implementing gamified leaderboards on Moodle.

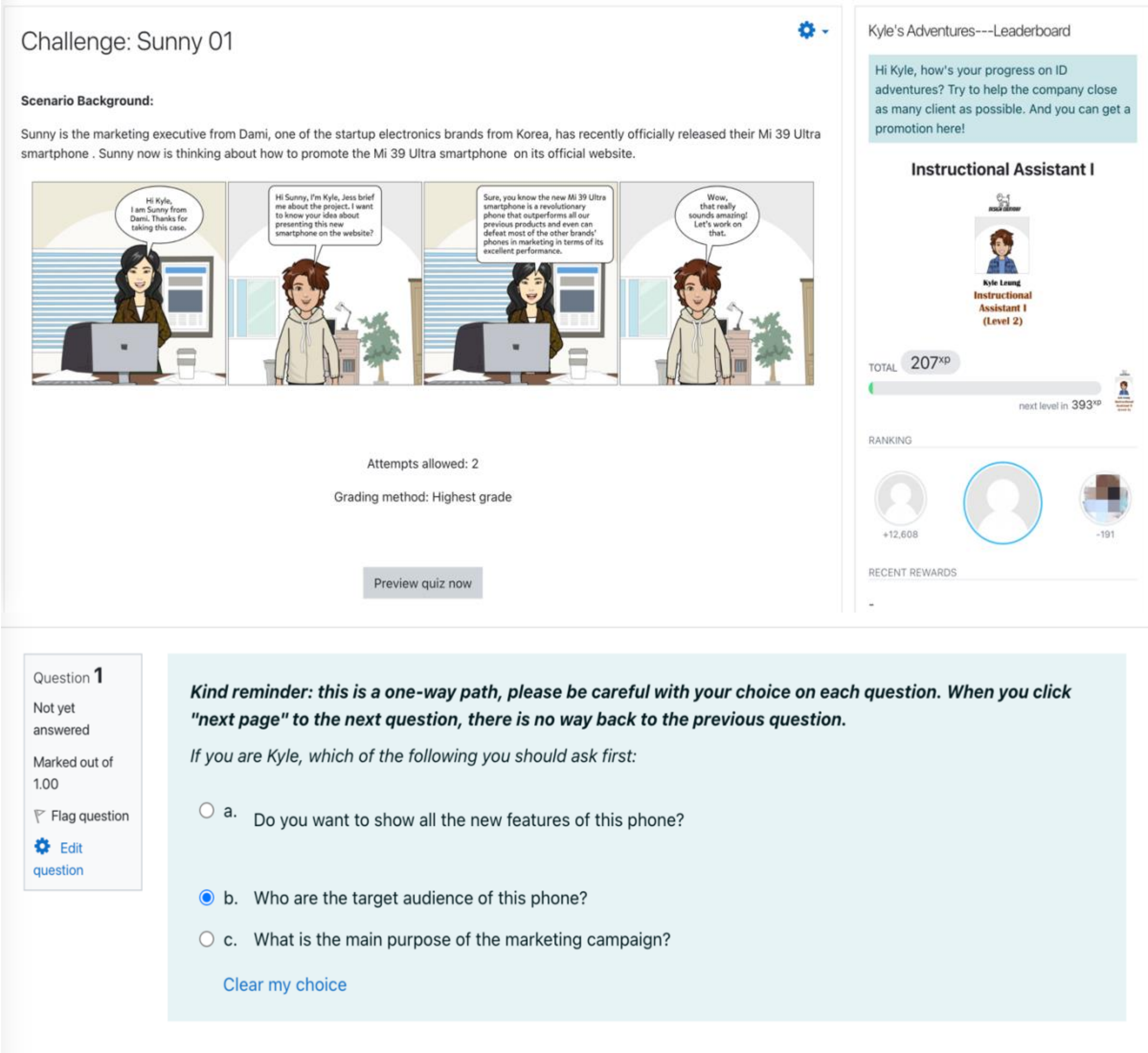
Points were awarded to individual students once the completion or performance criteria for the learning tasks were met. The learning tasks were based on real-world problems, such as designing a training course for newly hired employees. Completing these tasks helped the students develop their understanding and application of the knowledge and skills acquired. Completion-contingent points were awarded when a participant completed a task. For example, we awarded 1 point for each course viewing action. As the gamified course included 12 levels, we could award a maximum of 16,200 points. These levels consisted of a total of 150 tasks. The students' participation scores increased after completing 75 tasks to reach level 8, which were labeled as graded tasks. All additional tasks completed (tasks 76–150) did not affect their course grades. To trigger initial student interest in gamification, 10% of the course grade was awarded as a participation score. The students were required to complete 75 tasks and reach level 8 to achieve a full participation score. Appendix A indicates the number of points required for each level.

However, more performance-contingent than completion-contingent points were awarded. Performance-contingent points were awarded according to a participant's performance. We awarded 100 points for achieving a passing grade on a task in easy mode. We awarded 200 points if a student achieved a passing grade on a task in hard mode. For example, one task examined the students' course design skills in a real-world scenario, in which the application of appropriate teaching strategy were considered challenging (see Appendix B for details). Figure 1 shows a background scenario and a question for an individual task with the leaderboard main interface.

3.1.1. Using the GAFCC-F model to gamify the course

The two gamified courses in this study were designed using the GAFCC-F model, which stands for Goals, Access, Feedback, Challenges, Collaboration, and Fantasy (Bai et al., 2022). The GAFCC-F model was applied as follows: (1) Goals: The students were asked to complete course design tasks by playing the role of the main fictional character, Kyle, throughout the semester. The higher the level, the more challenging the task became. The ultimate goal was for Kyle to get promoted to company director (level 12). (2) Access: The students could access the hard mode tasks once they had successfully completed the easy mode tasks. (3) Feedback: The process of adding points for completing learning tasks was automated to award performance-contingent points for achieving a passing grade and completion-contingent points for work submission. Written feedback from the instructor was also provided on completion-contingent tasks. (4) Challenges: The course consisted of multiple individual and group tasks conducted throughout the semester. Completing easy mode tasks was a prerequisite for unlocking hard mode tasks. (5) Collaboration: The students formed study groups of three to four members during the first session of the course. Three group tasks were applied in both experiments, and each group had to complete several course design plans to conceptualize their final group project. (6) Fantasy: The endogenous fantasy setting required a close connection between the fantasy context and the learning content. The settings, plots, and Kyle's missions were related to the design of the e-learning course, making the tasks more meaningful and engaging. The students could easily self-identify themselves as Kyle in a virtual workspace and apply their instructional design skills to real-world scenarios for self-learning assessment.

Figure 1. A background scenario and a question for an individual task with the leaderboard main interface



3.2. Measures

The same data collection measures were used in the two experiments. Based on their final position in the leaderboards at the end of the semester, the students in the two leaderboard settings were divided into three ranking groups: top third, middle third, and bottom third. The independent variable was the type of leaderboard (i.e., absolute or relative). The dependent variables were the students' learning engagement, problem-based learning strategies, learning performance, and learning perceptions. We compared the four learning outcomes for the students of the same rank in the two leaderboard groups, such as students ranked in the top third of the absolute and relative leaderboard groups.

3.2.1. Coding problem-based learning strategies

We used ENA to conduct a comparative analysis of problem-based learning strategies across the two leaderboard conditions (Shaffer et al., 2016). Originally developed to model relationships among cognitive elements (Shaffer et al., 2009), ENA has since evolved into a flexible tool for visualizing and comparing connections between various activities, such as student learning behaviors captured in log files or chat interactions (Shaffer et al., 2016; Uzir et al., 2020). Unlike sequential analysis, which depends on large datasets, focuses on the order of events, and is often difficult to interpret, ENA can effectively capture, visualize, and statistically compare learning activity patterns even in small datasets (Csanadi et al., 2018). By generating content-focused summary statistics, ENA enables us to compare cognitive networks based on the strength of connections between specific elements, rather than structural measures such as network density, which only indicates the overall level of network connection without detailing specific relationships (Bowman et al., 2021). Additionally, compared with social network analysis, which emphasizes interpersonal relationships, ENA identifies behavioral patterns from trace data and applies statistical tests to compare these patterns under different conditions (Swiecki & Shaffer, 2020). This makes ENA particularly suited to the analysis of problem-based learning, where understanding the relationships between cognitive processes is crucial.

In our study, we applied ENA to examine how the students engaged in problem-based learning under two leaderboard conditions. The process involved three key steps. First, we collected trace data from student interactions, ensuring their accuracy through a rigorous coding process. Two independent coders achieved 98% inter-coder reliability, resolving all discrepancies to achieve 100% agreement. Second, we developed a coding framework aligned with our research questions, identifying four key processes: (a) *Analyze the given information*, which included actions such as viewing the background scenario and task attempts; (b) *Identify associated knowledge*, which involved students revisiting teaching slides to review relevant knowledge while working on the tasks; (c) *Apply learning knowledge to problems*, which included actions such as starting or submitting a task attempt; (d) *Reflect on the solution*, which included any review actions after a task attempt, such as reviewing the attempt or viewing a summary. Third, the finalized dataset was imported into ENA software (https://www.epistemicnetwork.org/) for analysis, allowing us to examine the relationships between these processes and compare patterns between the two leaderboard conditions.

3.2.2. Using trace data to measure learning engagement

In addition to the four key processes discussed above, we included two additional behavioral engagement data codes: *Average time taken to complete graded tasks* and *Number of non-graded tasks completed.* As explained, completing 75 tasks to reach level 8 affected the students' participation scores, which were labeled as graded tasks. Any additional tasks completed (tasks 76–150) that did not affect participation scores were labeled as non-graded task. These two behavioral data codes were used to assess the students' engagement in high-stakes and no-stakes tasks in the two leaderboard settings.

3.2.3. Using exams and coursework to measure learning performance

A pre-test and post-test approach was taken to assess the students' learning performance in the first and last sessions out of a total of 10 sessions. The pre-test aimed to examine the students' prior knowledge and consisted of two multiple-choice questions and five short essay questions. The post-test consisted of an individual course design prototype and a group online course design task. The maximum grade for each performance test was 100.

3.2.4. Using open-ended questions to measure learning perceptions

We conducted an open-ended survey after the semester to obtain the students' overall perceptions of the use of absolute and relative leaderboards. The survey questions were as follows: *How do you feel when your name and score are displayed on the leaderboard? How would you describe your learning experience using the leaderboard throughout the semester?* In our data analysis, we used the directed content analysis method described by Hickey and Kipping (1996), which provides a structured coding process allowing for the immediate application of pre-existing codes from previous research (Hsieh & Shannon, 2005). When the text did not fit the initial coding scheme, a new code was created. We also conducted a frequency analysis of emerging themes.

To enhance coding consistency, we identified several examples that clearly illustrated each theme. Multiple reviews of the data were conducted to ensure a thorough understanding of each theme (Creswell, 2012). All data were cross-coded by two coders to establish reliability. Additionally, the coding results were compared to assess inter-rater reliability, achieving an overall Cohen's κ of 0.92, where a value of 0.75 indicates substantial agreement (Viera & Garrett, 2005). Any discrepancies were carefully addressed through collaborative discussions between the two coders to reach an agreement.

3.3. Experiment 1: Absolute leaderboard

3.3.1. Participants

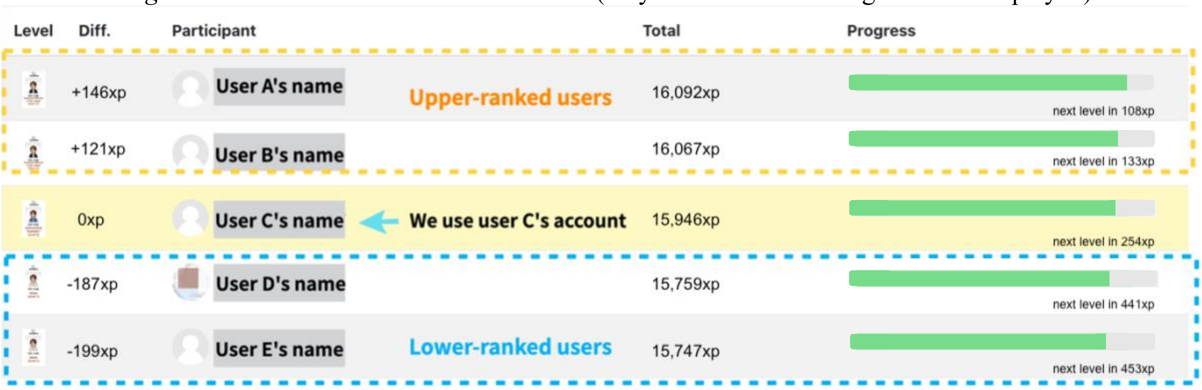
The absolute leaderboard setting included 24 postgraduate students (17 female students and 7 male students). Their ages ranged from 22 to 44 years (M = 28, SD = 6.22). Fourteen were from mainland China, four from Hong Kong SAR, one from the US, and one from Pakistan.

3.3.2. Absolute leaderboard setting

Figure 2 shows the absolute leaderboard interface on Moodle. The students in the class could access various information through their Moodle accounts, including (a) their rank in the leaderboard, (b) their current level, (c) their profile picture and name, (d) the names of other users, (e) total points collected, and (f) a progress bar indicating their proximity to the next level.

Figure 2. The absolute leaderboard interface (all participant names are displayed)

Figure 3. The relative leaderboard interface (only four immediate neighbors are displayed)



3.4. Experiment 2: Relative leaderboard

3.4.1. Participants

We enrolled 23 postgraduate participants in Experiment 2. They included 20 female students and 3 male students aged 21 to 47 years, with a mean age of 26.87 (SD = 6.91). Nineteen participants were from mainland China, three from Hong Kong SAR, and one from Indonesia.

3.4.2. Relative leaderboard setting

The relative leaderboard interface on Moodle displayed four neighboring participants (two immediately above and two immediately below the focal user). The left side of Figure 3 shows the levels, point differences, profile pictures, and names of the focal participant and their neighboring participants. The right side allows the focal participant (user C) to view the points collected by their four neighbors and their progress to the next level.

4. Results

4.1. Effects on student learning engagement

4.1.1. Top-ranked students' learning engagement

We first conducted normality tests for the six sets of behavioral data (see Table 1) in the two leaderboard groups. The data on *Average time taken to complete graded tasks* and *Number of non-graded tasks completed* in the two groups did not fit a normal distribution, while the remaining datasets did. Hence, we applied Mann–Whitney tests for *Average time taken to complete graded tasks* and *Number of non-graded tasks completed* and independent samples t-tests for the other four behavioral data comparisons.

The results showed that the top-ranked students in the relative leaderboard group significantly outperformed those in the absolute leaderboard group in terms of *Applying learning knowledge to problems*, $t(14) = 2.79$, $p = .014$, and completing graded tasks in less time on average with satisfactory quality, U = -2.37, $p = .018$ (see Table 1).

Table 1. Comparison of individual task completion behaviors of students across three ranking levels in the absolute and relative leaderboard groups

Individual task completion behaviors	Leaderboard type	Ranking position: Top third			Middle third			Bottom third		
		n	M(SD)/Mdn	Sig	n	M(SD)/Mdn	Sig	n	M(SD)/Mdn	Sig
Analyze the given information (c)	Relative	8	355.13 (18.51)	0.19	7	322.86 (38.44)	0.174	8	331^{*} (54.42)	0.013
	Absolute	8	388.75 (28.21)		8	291.38 (45.38)		8	261.13 (43.25)	
Identify associated knowledge (c)	Relative	8	145.5 (26.92)	0.79	7	113	0.3	8	134.75 (44.96)	0.058
	Absolute	8	149.25 (30.28)		8	109		8	98.25 (21.89)	
Apply learning knowledge to problems (c)	Relative	8	65.25^{*} (1.75)	0.014	7	61 (6.93)	0.092	8	64.5^{**}	0.009
	Absolute	8	61.13 (3.8)		8	54.25 (7.38)		8	51	
Reflect on the solution (c)	Relative	8	150.38 (18.99)	0.91	7	132.43 (15.76)	0.23	8	140.63^{**} (23.66)	0.004
	Absolute	8	151.38 (16.91)		8	122.25 (15.29)		8	106.88 (15.1)	
Average time taken to complete graded tasks (d)	Relative	8	28^{*}	0.018	7	28.14^{**} (1.77)	0.006	8	37.75 (6.45)	0.094
	Absolute	8	32.5		8	34.13 (4.55)		8	32.5 (5.18)	
Number of non-graded tasks completed (c)	Relative	8	15	0.064	7	15^{**} (0.01)	0.007	8	15^{***}	< 0.001
	Absolute	8	15		8	13.13 (1.55)		8	8	

Note. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. c = count; d = day; n = number of participants; M = mean; SD = standard deviation; Mdn = median; Sig = significance. Mann–Whitney tests were used for *Average time taken to complete graded tasks* and *Number of non-graded tasks completed*. The other four outcomes used independent samples t-tests.

4.1.2. Middle-ranked students' learning engagement

A normality test indicated that five behavioral datasets were normally distributed, while the dataset on *Identify associated knowledge* was not. Hence, we applied a non-parametric Mann–Whitney test to *Identify associated knowledge* and independent samples t-tests for the behavioral data comparisons of the other five datasets.
We found that the middle-ranked students in the relative leaderboard group completed their graded tasks in significantly less time on average with satisfactory quality, $t(13) = -3.26$, $p = .006$, and completed more non-graded tasks, $t(13) = 3.18$, $p = .007$, than those in the absolute leaderboard group (see Table 1).

4.1.3. Bottom-ranked students' learning engagement

The normality tests suggested that two datasets did not fit a normal distribution (i.e., *Apply learning knowledge to problems* and *Number of non-graded tasks completed*). Hence, we conducted Mann–Whitney tests for these two behavioral data comparisons and independent samples t-tests for the remaining four comparisons.

We found that the bottom-ranked students in the relative leaderboard group scored higher than those in the absolute leaderboard group in most areas of behavioral engagement. Specifically, the relative leaderboard group had more logs for *Analyze the given information* ($t(14) = 2.84$, $p = .013$), *Apply associated knowledge to problems* ($U = -2.63$, $p = .009$), and *Reflect on the solution* ($t(14) = 3.4$, $p = .004$), and the students *completed significantly more non-graded tasks* ($U = -3.26$, $p < .001$) (see Table 1).

Table 2. Connection coefficients of three ranking levels in ENA networks in the relative and absolute leaderboard groups

Connection	Relative_top-ranked	Absolute_top-ranked	Relative_middle-ranked	Absolute_middle-ranked	Relative_bottom-ranked	Absolute_bottom-ranked
Analyze the given information – Identify associated knowledge	0.283	0.277	0.269	0.271	0.279	0.283
Identify associated knowledge – Apply learning knowledge to problems	0.282	0.268	0.280	0.276	0.276	0.285
Apply learning knowledge to problems – Reflect on the solution	0.346	0.344	0.348	0.347	0.350	0.346
Analyze the given information – Reflect on the solution	0.653	0.659	0.667	0.663	0.656	0.666
Analyze the given information – Apply learning knowledge to problems	0.470	0.479	0.500	0.497	0.482	0.487
Identify associated knowledge – Reflect on the solution	0.255	0.247	0.180	0.201	0.232*	0.183

Note. $^*p < .05$. Relative_top= top-ranked participants in the relative leaderboard group; Absolute_top = top-ranked participants in the absolute leaderboard group.

4.2. Students' problem-based learning strategies

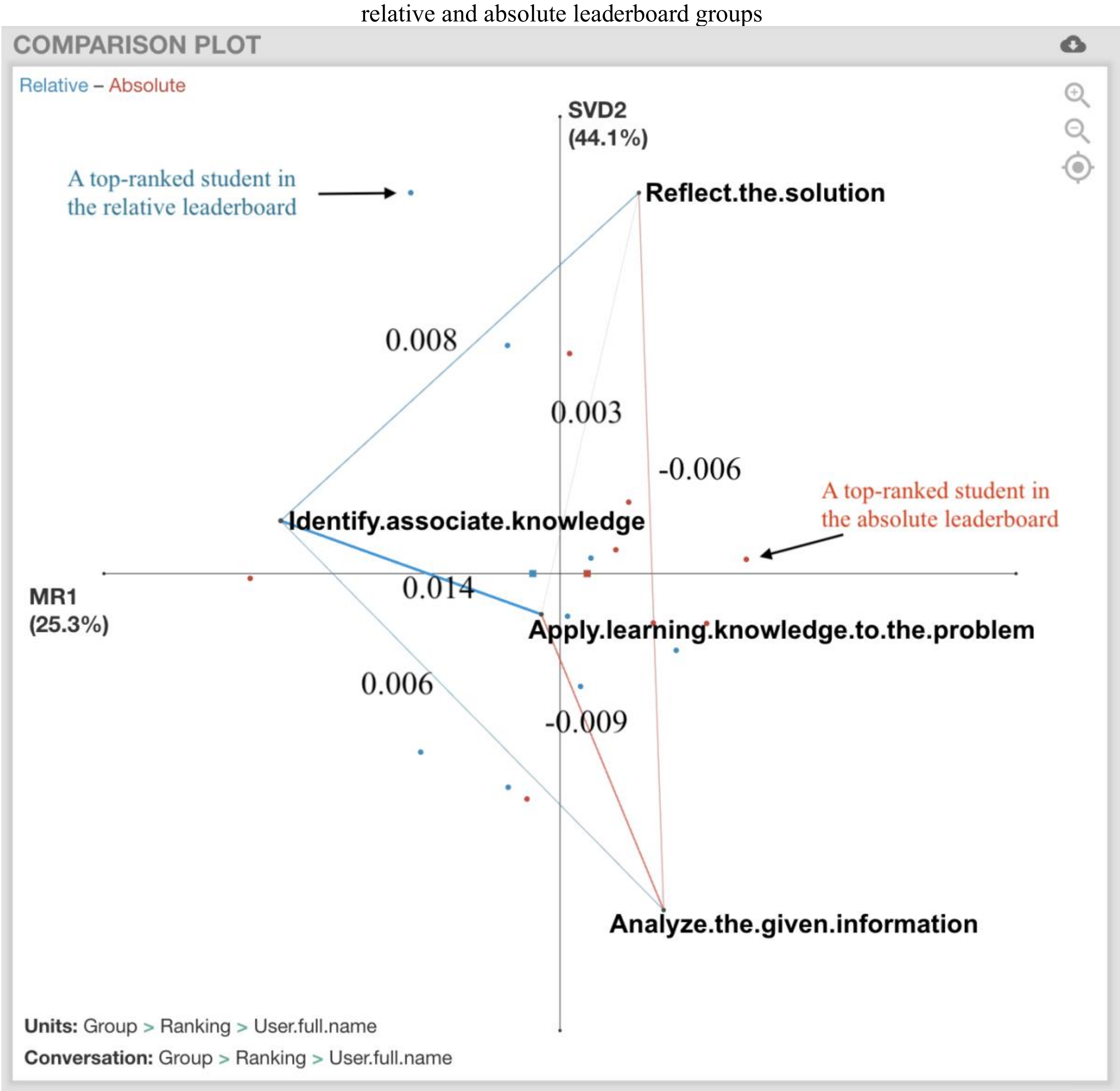
4.2.1. Top-ranked students' problem-based learning strategies

Figure 4 shows the subtracted epistemic network of the top-ranked students in the relative and absolute leaderboard groups in terms of differences in problem-based learning strategies. The largest difference appeared in the connection between *Identify associated knowledge* and *Apply learning knowledge to problems* (difference = 0.014), with the top-ranked students in the relative leaderboard group showing a stronger connection (0.282) than those in the absolute leaderboard group (0.268). However, none of the differences were significant, indicating a lack of significant distinction in problem-based learning strategies. See Table 2 for connection coefficients of three ranking levels in ENA networks in the relative and absolute leaderboard groups.

4.2.2. Middle-ranked students' problem-based learning strategies

Figure 5 shows the subtracted epistemic network of the middle-ranked students in the relative and absolute leaderboard groups in terms of differences in problem-based learning strategies. The largest difference appeared in the connection between *Identify associated knowledge* and *Reflect on the solution* (difference = -0.021), with the middle-ranked students in the absolute leaderboard group showing a stronger connection (0.201) than those in the relative leaderboard group (0.180). However, none of these differences were significant, indicating a lack of significant distinction in problem-based learning strategies.

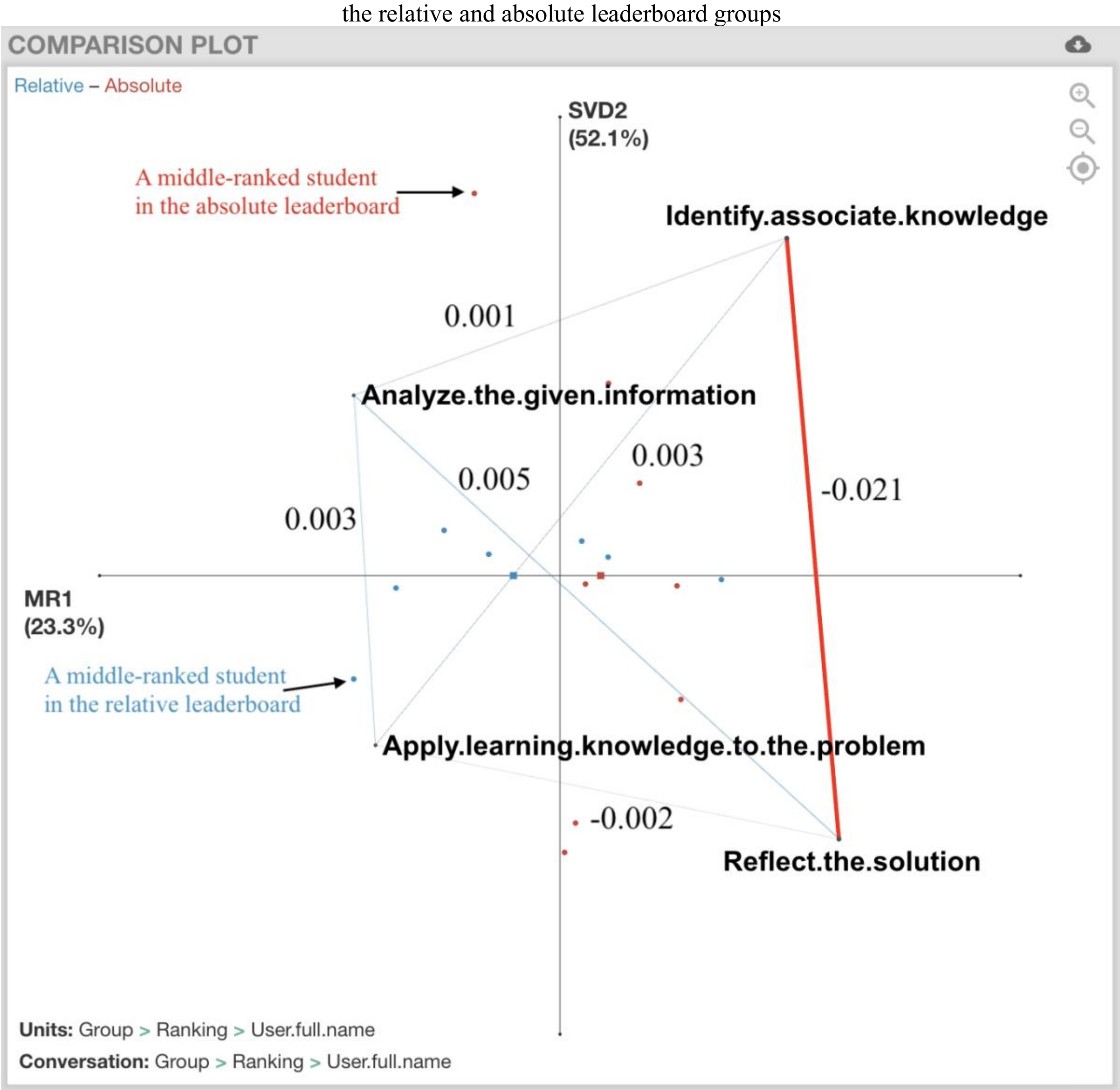
Figure 4. A subtracted epistemic network of problem-based learning strategies among top-ranked students in the relative and absolute leaderboard groups



Note. Red dot ● = a top-ranked student in the absolute leaderboard group; blue dot ● = a top-ranked student in the relative leaderboard group; red square ■ = the centroid (mean position of the projected points) of the top-third ranked students in the absolute leaderboard group; blue square ■ = the centroid (mean position of the projected points) of the top-third ranked students in the relative leaderboard group; MR = mean rotation; SVD = singular value decomposition.

Annotation: The subtracted epistemic network shows the differences between the top-ranked students in the absolute and relative leaderboard groups. The numerical values on the lines indicate the differences in connection coefficients between two codes. A thicker line indicates a greater difference. For example, a connection coefficient of 0.014 between the codes *Identify associated knowledge* and *Apply learning knowledge to problems* indicates the largest difference. The color of the line denotes the dominant group. The blue line with a connection coefficient of 0.014 indicates that the relative leaderboard group dominates the relationship between the codes *Identify associated knowledge* and *Apply learning knowledge to problems*.

Figure 5. A subtracted epistemic network of problem-based learning strategies among middle-ranked students in the relative and absolute leaderboard groups



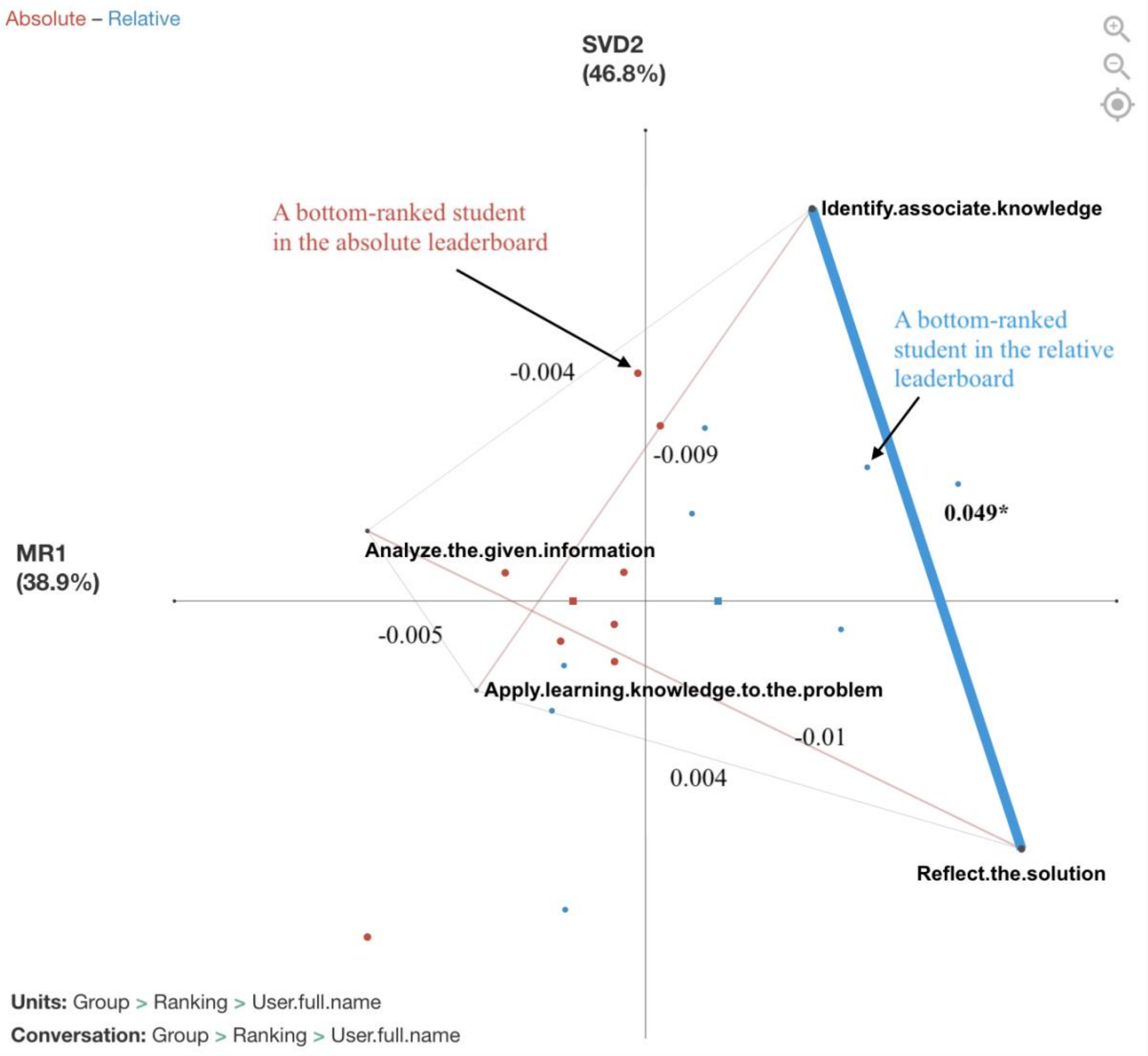
Note. Red dot ● = a middle-ranked student in the absolute leaderboard group; blue dot ● = a middle-ranked student in the relative leaderboard group; red square ■ = the centroid (mean position of the projected points) of the middle-third ranked students in the absolute leaderboard group; blue square ■ = the centroid (mean position of the projected points) of the middle-third ranked students in the relative leaderboard group; MR = mean rotation; SVD = singular value decomposition.

Annotation: The subtracted epistemic network shows the differences between the middle-ranked students in the absolute and relative leaderboard groups. The numerical values on the lines indicate the differences in connection coefficients between two codes. A thicker line indicates a greater difference. For example, a connection coefficient of -0.021 between the codes *Identify associated knowledge* and *Reflect on the solution* indicates the largest difference. The color of the line denotes the dominant group. The blue line with a connection coefficient of -0.021 indicates that the absolute leaderboard group dominates the relationship between the codes *Identify associated knowledge* and *Reflect on the solution*.

4.2.3. Bottom-ranked students' problem-based learning strategies

Figure 6 shows the subtracted epistemic network of the bottom-ranked students in the relative and absolute leaderboard groups in terms of differences in problem-based learning strategies. The most significant difference appeared in the connection between *Identify associated knowledge* and *Reflect on the solution* (difference = 0.049), with the bottom-ranked students in the relative leaderboard group showing a stronger connection (0.232) than those in the absolute leaderboard group (0.183). This difference was statistically significant ($p < .05$). This implies that the bottom-ranked students in the relative leaderboard group were more likely to review and reflect on their practices, prioritizing improved learning over rushing to complete tasks.

Figure 6. A subtracted epistemic network constructed using data on individual tasks completed from the log files of the bottom-ranked students in the relative and absolute leaderboard groups



Note. Red dot ● = a bottom-ranked student in the absolute leaderboard group; blue dot ● = a bottom-ranked student in the relative leaderboard group; red square ■ = the centroid (mean position of the projected points) of the bottom-third ranked students in the absolute leaderboard group; blue square ■ = the centroid (mean position of the projected points) of the bottom-third ranked students in the relative leaderboard group; MR = mean rotation; SVD = singular value decomposition.

Annotation: The subtracted epistemic network shows the differences between the bottom-ranked students in the absolute and relative leaderboard groups. The numerical values on the lines indicate the differences in connection coefficients between two codes. A thicker line indicates a greater difference. For example, a connection coefficient of 0.049 between the codes *Identify associated knowledge* and *Reflect on the solution* indicates the largest difference. The color of the line denotes the dominant group. The blue line with a connection coefficient of 0.049 indicates that the relative leaderboard group dominates the relationship between the codes *Identify associated knowledge* and *Reflect on the solution.*

4.3. Effects on students' learning performance

We compared the pre-test performance results and found non-significant differences between similar ranked students in the two leaderboard groups (e.g., top-ranked students in the relative and absolute leaderboard groups).

After confirming the normal distribution of all data for the learning performance measure, we found that the top-ranked students in the relative leaderboard group (M = 87.68, SD = 8.62) had significantly better learning performance in the post-test than the students in the absolute leaderboard group (M = 70.58, SD = 5.67), $t(14) = -4.69$, $p < .001$, effect size d = 2.34 (see Table 3).

The same result was found for the middle-ranked students. Indeed, the middle-ranked students in the relative leaderboard group (M = 84.65, SD = 5.15) also had significantly better learning performance in the post-test than the students in the absolute leaderboard group (M = 73.21, SD = 5.94), $t(13) = -3.95$, $p = .002$, effect size d = 2.06. However, there was no significant difference between the bottom-ranked students in the relative (M = 80.81, SD = 7.48) and absolute (M = 76.5, SD = 11.88) leaderboard groups, $t(14) = -.87$, $p = .4$, effect size d = 0.43.

Table 3. Comparison of learning performance among students at all ranking levels in the absolute and relative leaderboard groups

Stage	Leaderboard type	Top third			Middle third			Bottom third		
Pre-test		*n*	*M* (*SD*)	Effect size	*n*	*M* (*SD*)	Effect size	*n*	*M* (*SD*)	Effect size
	Relative	8	22.5 (9.93)	d = 1.06	7	14.81 (5.6)	d = 0.27	8	19.75 (15.18)	d = 0.03
	Absolute	8	11.43 (10.99)		8	17.67 (13.71)		8	20.17 (13.17)	
Post-test	Relative	8	87.68*** (8.62)	d = 2.34	7	84.65** (5.15)	d = 2.06	8	80.81 (7.48)	d = 0.43
	Absolute	8	70.58 (5.67)		8	73.21 (5.94)		8	76.5 (11.88)	

Note. ***$p < .001$, **$p < .01$. n = number of participants; M = mean; SD = standard deviation; d = Cohen's d.

4.4. Students' perceptions of relative and absolute leaderboards

Five themes emerged from the relative leaderboard group, which emphasized self-comparison and satisfaction: (1) *feel satisfied after recognizing my effort/progress* (n = 7, top-ranked students); (2) *work hard to keep the advantage* (n = 6, top-ranked students); (3) *a sense of uncertainty drives me to complete more tasks to level up* (n = 4, middle-ranked students); (4) *appreciate our diligence and progress* (n = 3, middle-ranked students); and (5) *feel happy to see my progress* (n = 4, bottom-ranked students).

Four themes emerged from the absolute leaderboard group, highlighting comparison concerns: (1) *motivate me to work harder* (n = 4, top-ranked students); (2) *increase peer comparison* (n = 3, middle-ranked students); (3) *struggle to level up* (n = 3, middle-ranked students); and (4) *feel upset and experience peer pressure after seeing my low ranking* (n = 3, bottom-ranked students).

The top-ranked students in the two leaderboard groups shared similar positive attitudes. Most of these students in the relative leaderboard group were satisfied and recognized their efforts and progress (7 out of 8): "*I am happy that my name is on the leaderboard because my hard work is paying off*" (Student 11). Six students said they worked hard to maintain their top ranking. As one student reported: "*I feel that it encouraged me to do more or at least maintain my top position during the course*" (Student 9). In the absolute leaderboard setting, half of the top-ranked students stated that the ranking scheme motivated them to work harder: "*It inspired me to work harder. When I was in the top five, I felt a strong sense of achievement*" (Student 2).

The middle-ranked students in the two leaderboard groups reported very different perceptions. Their relative rankings created a sense of uncertainty, which pushed them to complete more tasks to level up (4 out of 8): "*Because I couldn't see the ranking of the entire class, I only knew the scores of a few students close to me. A feeling of uncertainty drove me to try to see who else was in front of me. When I discovered that I had moved up the ranking, I was more motivated to complete more tasks*" (Student 27). Three students in the relative leaderboard group reported that their diligence and progress had been rewarded. However, the middle-ranked students in the absolute leaderboard group stressed the intensified level of peer comparison and struggled to level up. As one participant stated, "*It made me feel more competitive. I saw other classmates as my competitors*" (Student 14).

The bottom-ranked students in the two leaderboard groups had very different opinions about the ranking experience. Half of these students in the relative leaderboard group reported feeling happy after seeing their progress: "*I feel wonderful. I think it is a good way to motivate students to monitor their progress in this rewarding learning journey*" (Student 40). In contrast, most of these students in the absolute leaderboard group had a negative opinion or were indifferent. Three participants (out of 8) reported being upset and stressed: "*I am a little nervous because I am often ranked last. It makes me feel very incompetent*" (Student 32). Three participants were indifferent to their positions: "*I am fine with it. I don't really care about my position in the leaderboard*" (Student 38).

5. Discussion

5.1. The bottom-ranked students in the relative leaderboard group were more engaged than those in the absolute leaderboard group

Overall, the students in the relative leaderboard group demonstrated higher levels of engagement than those in the absolute leaderboard group, especially among the bottom-ranked students. Notably, the bottom-ranked students in the relative leaderboard group outperformed their counterparts in four of the six measures of behavioral engagement, including completing a greater number of non-graded learning tasks. Additionally, these students reported an overall positive attitude, finding satisfaction in comparing themselves with close competitors and recognizing their progress. Conversely, the bottom-ranked students in the absolute leaderboard group felt upset and stressed by their low ranking, experiencing significant peer pressure.

This discrepancy may be attributed to the higher peer pressure involved in the absolute leaderboard group, which may have induced anxiety and gradually reduced interest in competition (Albuquerque et al., 2017; Hanus & Fox, 2015). The transparency of information about the bottom-ranked students in the absolute leaderboard group exacerbated this issue. In contrast, the relative leaderboard design helped the students focus on the attainable goal of outperforming a limited number of competitors ranked above them. This fostered a supportive and motivating learning environment by reducing the stress associated with moving up the ranks. Previous studies have indicated that public displays of poor performance are demotivating (Leung, 2019) and can negatively impact self-esteem (Chan et al., 2018). Our finding aligns with the literature suggesting that bottom-ranked students are less likely to remain engaged in closing the gap with their top-ranked peers in an absolute leaderboard (Na & Han, 2023; Qiao et al., 2024). Upward comparisons in an absolute leaderboard can erode motivation to close the gap (Wang et al., 2024).

Moreover, the perceived effectiveness of absolute leaderboards may be influenced by the number of competitors. In our study, 24 students were part of the absolute leaderboard group. Lessel et al. (2022) suggested that the appeal of absolute leaderboards diminishes as the number of competitors increases, although the specific threshold for this decline remains unclear. This highlights the need for further exploration to determine the number of competitors at which this shift occurs (Lessel et al., 2022).

5.2. The top-ranked students were the most engaged regardless of the leaderboard type

Regardless of the leaderboard type, the top-ranked students exhibited the highest level of engagement compared with their peers in other ranks. These students completed graded tasks faster and successfully completed all non-graded tasks. Engagement levels, measured by six behaviors, tended to decrease from top-ranked students to bottom-ranked students and to middle-ranked students.

This observation aligns with the findings of Yang and Koenigstorfer (2025), who explored how different ranking positions affect users' motivation to engage in a physical activity. They discovered a curvilinear U-shaped effect: individuals at the top and bottom of the rankings were more motivated to complete tasks than those in the middle ranks. Specifically, the top-ranked individuals were driven by the desire to maintain their high status and enjoy the recognition that comes with it. In addition, the bottom-ranked individuals strove to improve their ranking and escape the negative connotations associated with being last. In contrast, the middle-ranked individuals might lack this sense of urgency or recognition, resulting in lower levels of motivation. These insights suggest that the extremes of the ranking spectrum (both high and low) serve as powerful motivators for student engagement and performance, while middle rankings require different strategies to bolster engagement.

5.3. The top-ranked students in the relative leaderboard group had better learning performance

In our analysis of descriptive statistics, we found that the post-test scores of all ranking cohorts in the relative leaderboard group were higher than those in the absolute leaderboard group. Specifically, the learning performance of the top- and middle-ranked students in the relative leaderboard group was significantly higher than that of their counterparts in the absolute leaderboard group. Notably, the top-ranked students in the relative leaderboard group achieved the best learning performance among all participants. An overall comparison of ranks and leaderboard types revealed that the top-ranked students in the relative leaderboard group were the most engaged (see Table 1), which probably contributed to their superior performance. However, there was no significant difference in terms of learning performance between the bottom-ranked students in the two leaderboard groups, although descriptively, the relative leaderboard group performed better. The marginally better performance of the bottom-ranked students in the relative leaderboard group could stem from the reduced stress of comparing themselves to peers in similar positions. This may create a slightly more supportive environment than for the students in the absolute leaderboard group, where the gap with the top performers was more pronounced.

5.4. Practical implications

Our findings lead to the following two practical implications. First, *use relative leaderboards to foster constructive competition*. Teachers might consider using relative leaderboards in their classrooms to foster a supportive and competitive learning environment. Our findings indicate that students, especially those at the bottom of the rankings, show higher engagement and a greater tendency to review and reflect on their learning when using relative leaderboards. This approach encourages students to focus on understanding and mastery rather than simply completing tasks, which can lead to improved learning strategies and overall engagement. Second, *limit the public display of the rankings of bottom-ranked students to enhance their learning motivation and engagement*. To avoid the negative impacts of peer pressure and anxiety, teachers should limit the public display of rankings, especially for bottom-ranked students. Our study shows that the absolute leaderboard group experienced higher levels of anxiety and reduced engagement than the relative leaderboard group.

6. Limitations and future research

This study has three main limitations. First, the use of the two leaderboard types was relatively short (10 weeks each), so the participants may have experienced a novelty effect. Extending the duration of the intervention could be beneficial in future studies; for instance, a one-year longitudinal study with the same set of participants would be ideal.

Second, the small sample sizes for the two groups (24 and 23 participants) limit the generalizability of our findings. Additionally, the modest statistical power (0.4) of our study reduced our ability to detect significant effects, making it more difficult to draw definitive conclusions. Therefore, caution should be exercised when interpreting the results of our study. Future studies should involve larger samples to enhance the robustness and generalizability of the results.

Third, the demographic details of the participants, such as gender imbalance and cultural homogeneity, warrant careful consideration in interpreting the findings. In our study, more than 70% of the participants in both experiments were women, which may have influenced the results. Research has suggested that women tend to be less competitive than men, partly due to less favorable beliefs about the outcomes of competition (Carpenter et al., 2018; Kesebir, 2020). This gender disparity raises questions about the generalizability of the findings to more gender-balanced or male-dominated samples.

Furthermore, the cultural composition of the sample, which was predominantly Chinese, may have introduced additional contextual influences. Cultural differences can influence how individuals perceive and respond to competition (Wu & Talhelm, 2023). Cultural psychology has shown that East Asian individuals (e.g., Chinese) tend to emphasize a collective self, while Americans often focus on an individual self (Heine, 2001).

Although collectivistic cultures are traditionally associated with a sense of harmony, recent evidence suggests that they may compete more intensely than individualistic cultures (Wu & Talhelm, 2023).

Given these demographic and cultural factors, the findings of this study should be interpreted with caution. Future research should incorporate more diverse participant pools, including varying gender compositions and cultural backgrounds, to better understand competitive motivation in learning. Such efforts would help enhance the generalizability of the findings to different cultural and demographic contexts.

7. Conclusion

In this study, we conducted a quasi-experiment to compare the effects of absolute and relative leaderboards on students' learning engagement, problem-based learning strategies, learning performance, and perceptions in a fully online course setting. The results showed that compared with the absolute leaderboard group, comparison with neighboring competitors in the relative leaderboard group led to a higher level of learning engagement and performance and encouraged constructive competitiveness that prompted the students to prioritize knowledge mastery over mere task completion. The students at all ranking levels in the relative leaderboard group reported a satisfied, motivated, and progress-oriented attitude. In contrast, the students at different ranking levels in the absolute leaderboard group showed different attitudes. Although the top-ranked students were motivated to work harder and were happy to see their achievements publicly displayed, the middle- and bottom-ranked students reported increased peer pressure as they progressed.

This study enhances our understanding of how leaderboard configurations influence student behavior and outcomes in online learning environments. The findings highlight the potential of relative leaderboards to create a supportive and motivating environment, particularly benefiting lower-ranked students by encouraging self-improvement and mastery. Future research should explore the long-term effects of relative leaderboards in different cultural contexts and use larger samples. Understanding how cultural factors shape student responses will provide deeper insights into the effectiveness of gamification strategies. Additionally, longitudinal studies could examine the sustained impact of these leaderboards on students' motivation, engagement, and academic performance, thereby providing a more comprehensive understanding of their potential in educational settings.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Availability of data and material

Data available upon request from the corresponding author due to ethical restrictions.

Acknowledgment

The research is supported by a grant from the Research Grants Council of Hong Kong (Project reference no: ECS 28600024). Large language Models (e.g., DeepSeek v3) was used to improve the clarity of some writings in this manuscript. After using this tool, the authors reviewed and edited the specific contents and take full responsibility for the accuracy of the contents. The manuscript was later sent for professional English proofreading for final editing.

References

Albuquerque, J., Bittencourt, I. I., Coelho, J. A. P. M., & Silva, A. P. (2017). Does gender stereotype threat in gamified educational environments cause anxiety? An experimental study. *Computers & Education*, *115*, 161–170. https://doi.org/10.1016/J.COMPEDU.2017.08.005

Bai, S., Hew, K. F., & Gonda, D. E. (2021). Examining effects of different leaderboards on students' learning performance, intrinsic motivation, and perception in gamified online learning setting. *2021 IEEE International Conference on Educational Technology (ICET)*, 36–41. https://doi.org/10.1109/ICET52293.2021.9563130

Bai, S., Hew, K. F., Gonda, D. E., Huang, B., & Liang, X. (2022). Incorporating fantasy into gamification promotes student learning and quality of online interaction. *International Journal of Educational Technology in Higher Education*, *19*(1), 29. https://doi.org/10.1186/s41239-022-00335-9

Bai, S., & Liu, Y. (2024). Comparing effects of adaptive gamification and one-size-fits-all gamification on students' task completion process and learning performance. *The 32nd Interactional Conference on Computers in Education*, 1–10. https://doi.org/10.58459/icce.2024.4892

Balci, S., Secaur, J. M., & Morris, B. J. (2022). Comparing the effectiveness of badges and leaderboards on academic performance and motivation of students in fully versus partially gamified online physics classes. *Education and Information Technologies*, *27*(6), 8669–8704. https://doi.org/10.1007/s10639-022-10983-z

Bennani, S., Maalel, A., & Ben Ghezala, H. (2022). Adaptive gamification in E-learning: A literature review and future challenges. *Computer Applications in Engineering Education*, *30*(2), 628–642. https://doi.org/10.1002/cae.22477

Bowman, D., Swiecki, Z., Cai, Z., Wang, Y., Eagan, B., Linderoth, J., & Shaffer, D. W. (2021). The mathematical foundations of epistemic network analysis. In: Ruis, A.R., Lee, S.B. (eds) Advances in quantitative ethnography. ICQE 2021. *Communications in Computer and Information Science*, vol 1312. Springer, Cham. https://doi.org/10.1007/978-3-030-67788-6_7

Carpenter, J., Frank, R., & Huet-Vaughn, E. (2018). Gender differences in interpersonal and intrapersonal competitive behavior. *Journal of Behavioral and Experimental Economics*, *77*, 170–176. https://doi.org/10.1016/j.socec.2018.10.003

Chan, E., Nah, F. F. H., Liu, Q., & Lu, Z. (2018). Effect of gamification on intrinsic motivation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10923 LNCS*, 445–454. https://doi.org/10.1007/978-3-319-91716-0_35/TABLES/3

Cigdem, H., Ozturk, M., Karabacak, Y., Atik, N., Gürkan, S., & Aldemir, M. H. (2024). Unlocking student engagement and achievement: The impact of leaderboard gamification in online formative assessment for engineering education. *Education and Information Technologies*, *29*(18), 24835–24860. https://doi.org/10.1007/s10639-024-12845-2

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.

Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *13*(4), 419–438. https://doi.org/10.1007/S11412-018-9292-Z/FIGURES/3

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

Do, N., Jin, T., Priest, R., Meredith, L. N., & Landers, R. N. (2024). A longitudinal quasi-experiment of leaderboard effectiveness on learner behaviors and course performance. *Learning and Individual Differences*, *116*, 102572. https://doi.org/10.1016/j.lindif.2024.102572

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Gao, F. (2024). Advancing gamification research and practice with three underexplored ideas in self-determination theory. *TechTrends*, *68*(4), 661–671. https://doi.org/10.1007/s11528-024-00968-9

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *8*(6), 634–650. https://doi.org/10.1177/1745691613504114

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*(2), 177–197. https://doi.org/10.1037/bul0000127

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

Heine, S. J. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of personality, 69*(6), 881-905. https://doi.org/10.1111/1467-6494.696168

Hew, K. F., & Lee, C. K. (2019). Using digital badges and leaderboards in primary school math lessons: Beneficial or merely new wine in old bottles? *In Shaping the Future of Education, Communication and Technology: Selected Papers from the HKAECT 2019 International Conference* (pp. 71-90). Springer.

Hickey, G., & Kipping, C. (1996). A multi-stage approach to the coding of data from open-ended questions. *Nurse Researcher*, *4*(1), 81–91. https://doi.org/10.7748/nr.4.1.81.s9

Houshmand, K. (2015). *How to avoid the destructive impacts of competition at school*. Huffington Post. https://www.huffpost.com/archive/ca/entry/how-to-avoid-the-destructive-impacts-of-competition-at-school_b_8156376

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Kesebir, S. (2020). *How women and men view competition differently*. American Association for Physician Leadership. https://www.physicianleaders.org/news/how-women-and-men-view-competition-differently

Landers, R. N., & Landers, A. K. (2014). An empirical test of the theory of gamified learning. *Simulation & Gaming,* 45(6), 769–785. https://doi.org/10.1177/1046878114563662

Lavoué, E., Monterrat, B., Desmarais, M., & George, S. (2019). Adaptive gamification for learning environments. *IEEE Transactions on Learning Technologies*, 12(1), 16–28. https://doi.org/10.1109/TLT.2018.2823710

Lee, B. W., Lee, M. K., & Fredendall, L. (2022). Response to relative performance feedback in simulation games. *The International Journal of Management Education*, *20*(3), 100698. https://doi.org/10.1016/j.ijme.2022.100698

Lessel, P., Altmeyer, M., Schubhan, M., Gudea, D., & Krüger, A. (2022). "Absolute or relative?" – Exploring the choice between leaderboard types in an image tagging task. *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play*, 63–69. https://doi.org/10.1145/3505270.3558346

Leung, W. (2019). How do one's peers on a leaderboard affect oneself? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–11. https://doi.org/10.1145/3290605.3300397

Li, C., Liang, L., Fryer, L. K., & Shum, A. (2024). The use of leaderboards in education: A systematic review of empirical evidence in higher education. *Journal of Computer Assisted Learning*, *40*(6), 3406–3442. https://doi.org/10.1111/jcal.13077

Na, K., & Han, K. (2023). How leaderboard positions shape our motivation: The impact of competence satisfaction and competence frustration on motivation in a gamified crowdsourcing task. *Internet Research*, *33*(7), 1–18. https://doi.org/10.1108/INTR-12-2021-0897

Nadeem, M., Oroszlanyova, M., & Farag, W. (2023). Effect of digital game-based learning on student engagement and motivation. *Computers*, *12*(9), 177. https://doi.org/10.3390/computers12090177

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/https://doi.org/10.1002/cae.12116

Ortiz-Rojas, M., Chiluiza, K., Valcke, M., & Bolanos-Mendoza, C. (2025). How gamification boosts learning in STEM higher education: A mixed methods study. *International Journal of STEM Education*, *12*(1), 1. https://doi.org/10.1186/s40594-024-00521-3

Park, S., & Kim, S. (2021). Is sustainable online learning possible with gamification? —The effect of gamified online learning on student learning. *Sustainability*, *13*(8), 4267. https://doi.org/10.3390/su13084267

Philpott, A., & Son, J.-B. (2022). Leaderboards in an EFL course: Student performance and motivation. *Computers & Education*, *190*, 104605. https://doi.org/10.1016/J.COMPEDU.2022.104605

Pi, Z., Yu, Q., Zhang, Y., Li, Y., Chen, H., & Yang, J. (2024). Presenting points or rank: The impacts of leaderboard elements on English vocabulary learning through video lectures. *Journal of Computer Assisted Learning*, *40*(1), 104–117. https://doi.org/10.1111/jcal.12871

Qiao, S., Yeung, S. S., Shen, X. A., Leung, K. L., Ng, T. K., & Chu, S. K. W. (2024). How competitive, cooperative, and collaborative gamification impacts student learning and engagement. *Language Learning & Technology*, *28*(1), 1–19. https://hdl.handle.net/10125/73546

Reyssier, S. D., Serna, A., Hallifax, S., Marty, J.-C., Simonian, S., & Lavoué, E. (2024). How does adaptive gamification impact different types of student motivation over time? *Interactive Learning Environments*, *32*(10), 6043–6062. https://doi.org/10.1080/10494820.2023.2248220

Schöbel, S. M., Janson, A., & Leimeister, J. M. (2023). Gamifying online training in management education to support emotional engagement and problem-solving skills. *Journal of Management Education*, *47*(2), 166–203. https://doi.org/10.1177/10525629221123287

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, *3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-Century assessment of learning. *International Journal of Learning and Media*, *1*(2), 33–53. https://doi.org/10.1162/ijlm.2009.0013

Sheridan, S., & Williams, P. (2011). Developing individual goals, shared goals, and the goals of others: Dimensions of constructive competition in learning contexts. *Scandinavian Journal of Educational Research*, *55*(2), 145–164. https://doi.org/10.1080/00313831.2011.554694

Sinnott, M., & Xia, L. A. (2020). A review of the Moodle gamification plugin "Level Up": Using a Moodle plugin to gamify learning of academic vocabulary. *International Journal of Computer-Assisted Language Learning and Teaching*, *10*(3), 89–95. https://doi.org/10.4018/IJCALLT.2020070107

Slamet, T., Brush, T., & Kwon, K. (2024). The effects of competition in gamified online discussions on learners' behavioral and cognitive engagement. *Technology, Knowledge and Learning*, 1–27. https://doi.org/10.1007/s10758-024-09740-7

Swiecki, Z., & Shaffer, D. W. (2020). iSENS: An integrated approach to combining epistemic and social network analyses. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 305–313. https://doi.org/10.1145/3375462.3375505

Uzir, N. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L.-A., & Fudge, A. (2020). Analytics of time management and learning strategies for effective online learning in blended environments. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 392–401. https://doi.org/10.1145/3375462.3375493

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5), 360–363.

Wang, Y. F., Hsu, Y. F., Fang, K. T., & Kuo, L. T. (2024). Gamification in medical education: Identifying and prioritizing key elements through Delphi method. *Medical Education Online, 29*(1), 2302231. https://doi.org/10.1080/10872981.2024.2302231

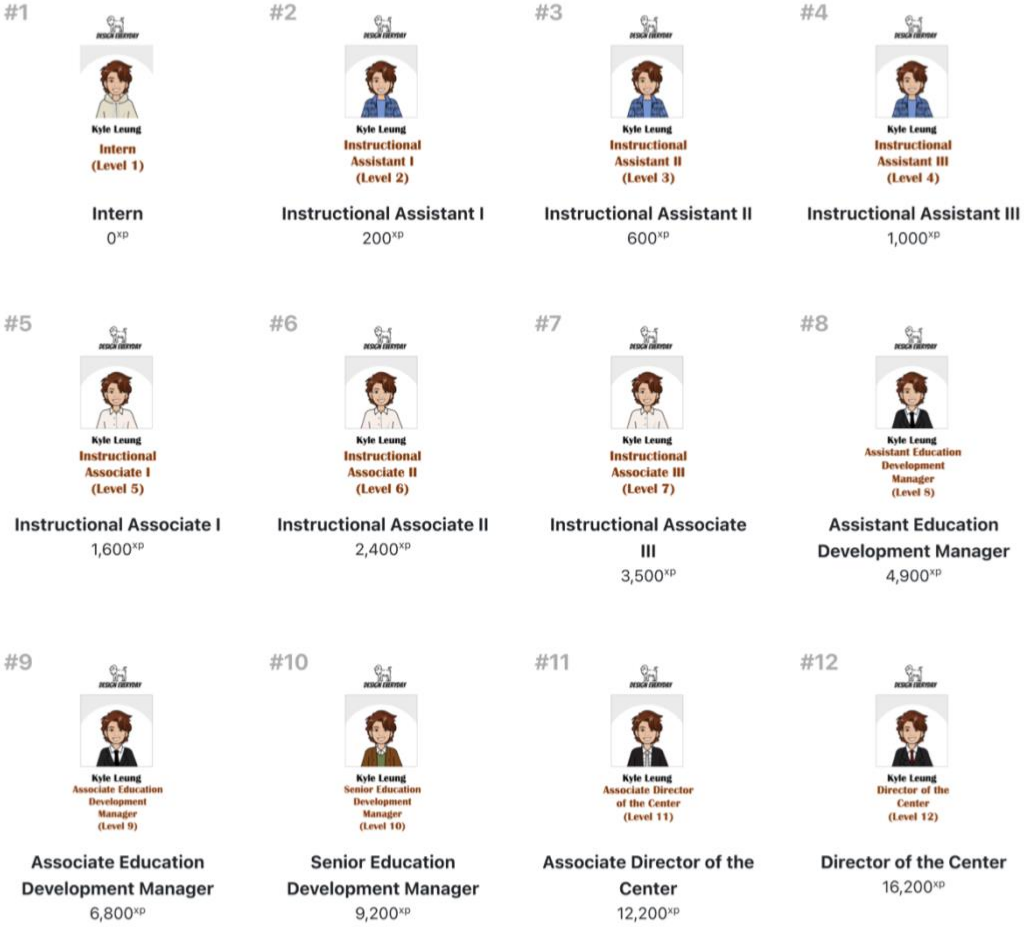
Wu, K., & Talhelm, T. (2023). Hide a dagger behind a smile: A review of how collectivistic cultures compete more than individualistic cultures. In S. M. Garcia, A. Tor & A. J. Elliot (Eds.), *The Oxford Handbook of the Psychology of Competition* (pp. 611–642). Oxford University Press.

Yang, Y., & Koenigstorfer, J. (2025). The curvilinear effects of relative positions in smartphone app leaderboards on physical activity. *Computers in Human Behavior,* 165, 108532. https://doi.org/10.1016/j.chb.2024.108532

Yang, Y., Goh, K. Y., Teo, H. H., & Tan, S. S. (2023). Compete with me? The impact of online gamified competition on exercise behavior. *Journal of the Association for Information Systems, 24*(3), 912–935. 10.17705/1jais.00806

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. Sebastopol, CA: O'Reilly Media.

Appendix A. Mapping 12 levels with points



Appendix B. An example of a task

> *Jenny Lo is a Human Resource executive at Pizza House. One day she calls you and asks about her company's new frontline employees' training design. You need to design a two-hour onboarding training for Pizza House frontline staff with two learning objectives: (a) to teach new employees about the various ingredients of pizza toppings; and (b) to create some materials that will help them to identify the different pizza ingredients.*

The purpose of this task was to evaluate the students' ability to apply the Analyse, Design, Develop, Implement, and Evaluate (ADDIE) instructional design model in course design. The ADDIE model is a systematic instructional design framework that includes five phases: Analyse the training needs, Design the training program, Develop the training materials, Implement the training session, and Evaluate its effectiveness. This exercise aimed to assess students' proficiency in systematically applying each phase of the ADDIE model to create an effective instructional design.