Cooperative Gamification for Introverted Learners: A Randomised Trial in Secondary STEM Classrooms

Emily H. Chen[1, *]

[1]University of British Columbia, Canada
E-mail: emily.h.chen@mock.iccmets.test

Abstract.

Extending our findings on leaderboard-induced demotivation, this randomised controlled trial in 24 secondary classrooms compares cooperative team-based gamification with individual leaderboards. Gamification is widely adopted in secondary STEM education, yet competitive mechanics such as public leaderboards have been shown to reduce motivation among learners who prefer low-stimulation, reflective environments. This study tested whether a cooperative alternative, in which points are earned collectively by small teams and progress is displayed at the team rather than the individual level, produces better outcomes for introverted learners without disadvantaging their extraverted peers. Twenty-four Grade 9 and 10 science and mathematics classrooms in British Columbia (N = 612 students) were randomly assigned to a cooperative team-based gamification condition or an individual leaderboard condition for a ten-week instructional unit. Trait introversion was measured at baseline, and behavioural engagement, task persistence and unit achievement were measured at baseline and after the intervention. Multilevel models with students nested in classrooms showed a significant condition by introversion interaction. Introverted learners in the cooperative condition showed significantly higher engagement and persistence than introverted learners in the leaderboard condition, with effect sizes of 0.48 and 0.41 respectively, while extraverted learners performed comparably across conditions. Achievement gains were modest and favoured the cooperative condition overall. The findings suggest that cooperative gamification offers a design that is inclusive of temperament differences and that leaderboard mechanics should be treated as optional rather than default in STEM classrooms.

Keywords: gamification, introversion, cooperative learning, STEM education, randomised controlled trial

**Corresponding author: emily.h.chen@mock.iccmets.test*

1. Introduction

Gamification, the application of game design elements such as points, badges and leaderboards to non-game contexts, has become a common feature of secondary science and mathematics instruction. Its appeal rests on the promise that game-like feedback will sustain effort on tasks that students might otherwise find tedious. The evidence base, however, is uneven. Meta-analyses report small positive average effects on motivation and learning, alongside considerable heterogeneity that depends on which mechanics are used and for whom (Sailer & Homner, 2020). Among the most widely used mechanics, the public leaderboard is also the most contested, because it introduces social comparison and competition into the classroom in a highly visible form.

In our earlier study (Chen & Cheungsirakulvit, 2025) we investigated why leaderboards demotivate introverted learners in STEM classrooms. Drawing on classroom observation and survey data, that work found that students high in trait introversion experienced public ranking as a source of evaluative threat rather than as a challenge, reported reduced

willingness to attempt difficult problems and disengaged from gamified activities at a higher rate than their extraverted peers. The study concluded that leaderboard mechanics impose a temperament-related cost that is not visible in class-level averages, and it called for experimental evaluation of alternative designs.

The present paper responds to that call. It reports a randomised controlled trial in 24 secondary classrooms comparing cooperative team-based gamification, in which points are earned collectively and displayed at the team level, with individual leaderboards. The primary question is whether introverted learners in the cooperative condition show higher engagement and persistence than introverted learners in the leaderboard condition. A secondary question is whether the cooperative design carries any cost for extraverted learners, who might be expected to thrive under competition. A tertiary question concerns achievement, since motivational benefits are of limited value if they do not translate into learning.

The study contributes to the gamification literature by moving from the description of a problem to the experimental test of a solution, and to the broader literature on individual differences in learning by examining temperament as a moderator of instructional design. It also contributes to practice by offering a design that teachers can implement with existing classroom tools. The paper reviews relevant theory and evidence, describes the trial design and measures, presents the results of multilevel analyses, discusses their implications and concludes with limitations.

2. Literature Review

2.1 Gamification and motivation

Deterding et al. (2011) provided the now standard definition of gamification as the use of game design elements in non-game contexts and distinguished it from full games and from playful design. Research on its educational effects has drawn heavily on self-determination theory (Deci & Ryan, 2000), which holds that intrinsic motivation is supported when activities satisfy needs for autonomy, competence and relatedness. Points and badges can support competence by providing feedback, but leaderboards may undermine autonomy by imposing external evaluation and may frustrate relatedness by positioning classmates as rivals. Hanus and Fox (2015) found in a semester-long study that a gamified course with leaderboards and badges reduced intrinsic motivation and final grades relative to a non-gamified course, while Sailer and Homner (2020) concluded from meta-analysis that effects on motivation are small and that competitive designs are more variable in their effects than cooperative ones.

2.2 Competition, cooperation and social comparison

Social comparison theory (Festinger, 1954) holds that individuals evaluate their abilities by comparison with others, and that comparison is most consequential when it is public and when the comparison target is similar. Leaderboards make comparison public and continuous. Cooperative learning research, by contrast, has established that structures in which students depend on one another for success promote both achievement and positive peer relations (Johnson & Johnson, 2009). Slavin (2014) argues that the key ingredients are group goals and individual accountability, so that team members have reason to support one another's learning. Cooperative gamification applies these principles to game mechanics: points are earned by the team, teams rather than individuals are displayed, and individual contribution is tracked privately. Morschheuser et al. (2019) found in a field experiment that cooperative gamification produced higher enjoyment and more prosocial behaviour than competitive gamification in a crowdsourcing setting, although evidence in classrooms remains sparse.

2.3 Introversion as a moderator

Introversion is a stable temperament dimension associated with preference for lower levels of external stimulation, reflection before action and reduced reward sensitivity to social recognition (Eysenck, 1967). Cain (2012) drew attention to the ways in which schools and workplaces systematically favour extraverted behaviour, including through group activities and public performance. In educational research, introverted students have been found to prefer individual or small-group work, to participate less in whole-class discussion and to experience public evaluation as more aversive. Our earlier work (Chen & Cheungsirakulvit, 2025) extended these findings to gamified STEM classrooms, showing that leaderboards produced a measurable engagement gap by temperament. The present trial tests two hypotheses derived from this literature. H1 states that there is a condition by introversion interaction such that introverted learners show higher engagement and persistence under cooperative gamification than under individual leaderboards. H2 states that extraverted learners show no significant difference in engagement and persistence between conditions.

3. Methodology

3.1 Design and participants

A cluster-randomised controlled trial was conducted in which classrooms were the unit of randomisation. Twenty-four Grade 9 and 10 science and mathematics classrooms in eight public secondary schools in British Columbia participated during a ten-week instructional unit in the 2025–2026 school year. Classrooms were stratified by subject and grade and randomly assigned within strata to the cooperative team-based gamification condition (12 classrooms) or the individual leaderboard condition (12 classrooms). Of 641 enrolled students, 612 provided consent and complete baseline data and form the analytic sample; attrition at post-test was 4.1% and did not differ by condition. Table 1 presents the profile of the sample by condition. Ethical approval was granted by the University of British Columbia and the participating school districts, and parental consent and student assent were obtained.

Table 1: Sample profile by condition (N = 612)

Characteristic	Cooperative (n = 308)	Leaderboard (n = 304)	Total
Classrooms	12	12	24
Grade 9 students	162	155	317
Grade 10 students	146	149	295
Female students	154	150	304
Subject: science	158	153	311
Subject: mathematics	150	151	301
Introverted (above median)	152	154	306
Mean baseline engagement (SD)	3.41 (0.72)	3.38 (0.75)	3.40 (0.73)

Source: Authors' trial data (2026)

3.2 Intervention conditions

Both conditions used the same learning platform, the same curriculum unit, the same weekly problem sets and the same point values for completed tasks. In the leaderboard condition, each student's cumulative points were displayed on a classroom leaderboard visible on the platform and projected at the start of each lesson, ranking all students. In the cooperative condition, students were assigned to teams of four, points earned by members accumulated to the team, and a team-level display showed each team's progress toward a shared goal rather than a ranking against other teams. Individual contributions were visible

privately to the student and teacher. Teachers received a two-hour training session on their assigned condition and implementation was monitored through platform logs and two observation visits per classroom; fidelity was rated acceptable in all 24 classrooms.

3.3 Measures

Trait introversion was measured at baseline with the ten-item extraversion subscale of the Big Five Inventory, reverse-scored so that higher values indicate greater introversion ($\alpha = 0.83$). For descriptive reporting, students above the sample median are described as introverted, although the analyses treat introversion as a continuous variable. Behavioural engagement was measured with a six-item scale adapted from Fredricks et al. (2004) covering attention, effort and participation ($\alpha = 0.86$ at baseline, 0.88 at post-test), supplemented by platform logs of task attempts. Task persistence was measured as the proportion of difficult problems, defined as those in the top quartile of difficulty, that a student attempted at least twice after an initial incorrect response, drawn from platform logs. Achievement was measured with a common end-of-unit test developed with the participating teachers (30 items, KR-20 = 0.81), with a parallel form administered at baseline.

3.4 Analysis

Because students were nested in classrooms, outcomes were analysed with two-level multilevel models (Raudenbush & Bryk, 2002) in which post-test scores were regressed on condition, introversion, their interaction and the baseline measure of the outcome, with random intercepts for classrooms. Interaction effects were probed by estimating simple effects of condition at one standard deviation above and below the mean of introversion, and standardised effect sizes were computed relative to the pooled baseline standard deviation. Intraclass correlations at baseline ranged from 0.06 to 0.09. Missing post-test data were handled by multiple imputation.

4. Results

4.1 Descriptive statistics

Baseline equivalence was satisfactory. Conditions did not differ significantly on introversion, baseline engagement, baseline persistence or baseline achievement, and the proportion of introverted students was almost identical across conditions. Table 2 reports post-test means for engagement, persistence and achievement by condition and introversion group. Two patterns are visible in the descriptive data. Among introverted students, all three outcomes were higher in the cooperative condition than in the leaderboard condition. Among extraverted students, engagement and persistence were very similar across conditions, while achievement was slightly higher in the cooperative condition.

Table 2: Post-test means (SD) by condition and introversion group

Outcome	Introverted, cooperative (n = 152)	Introverted, leaderboard (n = 154)	Extraverted, cooperative (n = 156)	Extraverted, leaderboard (n = 150)
Engagement (1–5)	3.78 (0.68)	3.42 (0.74)	3.74 (0.70)	3.71 (0.71)
Persistence (proportion)	0.61 (0.21)	0.52 (0.23)	0.60 (0.22)	0.59 (0.22)
Task attempts (mean per week)	14.2 (4.1)	12.1 (4.5)	14.0 (4.3)	13.8 (4.4)
Achievement (0–	20.4 (5.1)	19.1 (5.4)	20.6 (5.0)	20.1 (5.2)

30)						

Source: Authors' trial data (2026)

4.2 Hypothesis testing

Table 3 presents the multilevel model estimates. For engagement, the condition by introversion interaction was significant (b = 0.19, p = 0.002), supporting H1. Simple effects showed that at one standard deviation above the mean of introversion, the cooperative condition was associated with higher post-test engagement than the leaderboard condition (d = 0.48, p < 0.001), while at one standard deviation below the mean the difference was small and non-significant (d = 0.05, p = 0.61), supporting H2. The pattern for persistence was the same: the interaction was significant (b = 0.05, p = 0.008), the cooperative advantage among introverted students was moderate (d = 0.41, p < 0.001) and the difference among extraverted students was negligible (d = 0.04, p = 0.69). For achievement, the main effect of condition was small but significant (d = 0.16, p = 0.042) in favour of the cooperative condition, and the interaction did not reach significance (b = 0.31, p = 0.11), although the descriptive pattern indicated that the achievement advantage was concentrated among introverted students.

Table 3: Multilevel model estimates for post-test outcomes (N = 612, 24 classrooms)

Predictor	**Engagement b (SE)**	**p**	**Persistence b (SE)**	**p**	**Achievement b (SE)**	**p**
Baseline outcome	0.54 (0.04)	< 0.001	0.47 (0.04)	< 0.001	0.62 (0.03)	< 0.001
Condition (cooperative = 1)	0.18 (0.06)	0.006	0.04 (0.02)	0.031	0.78 (0.38)	0.042
Introversion (centred)	−0.14 (0.04)	< 0.001	−0.03 (0.01)	0.012	−0.21 (0.24)	0.38
Condition × introversion	0.19 (0.06)	0.002	0.05 (0.02)	0.008	0.31 (0.19)	0.11
Classroom variance	0.03		0.004		1.12	
Residual variance	0.28		0.031		14.6	

Source: Authors' trial data (2026)

4.3 Supplementary analyses

Platform log data corroborated the self-report findings. Introverted students in the leaderboard condition made fewer weekly task attempts than any other group and showed a declining trajectory of attempts across the ten weeks, whereas introverted students in the cooperative condition maintained attempts at a level comparable to extraverted students. Analysis by subject showed that the interaction was somewhat stronger in mathematics than in science, although the difference between subjects was not significant. No significant differences were found by gender or grade. Teachers in the cooperative condition reported in post-intervention interviews that quieter students were more likely to ask teammates for help and to contribute to team problem solving than they had observed in prior gamified units, while teachers in the leaderboard condition reported that a subset of students had stopped attending to the leaderboard by mid-unit.

A manipulation check confirmed that students perceived the conditions as intended. At post-test, students in the leaderboard condition reported significantly higher awareness of their standing relative to classmates than students in the cooperative condition, and students in the cooperative condition reported significantly higher perceived interdependence with peers. Perceived social comparison was, in turn, negatively associated with post-test engagement

among students high in introversion but not among those low in introversion, which is consistent with the proposed mechanism. Sensitivity analyses that excluded the two classrooms with the lowest fidelity ratings, and that used complete cases rather than imputed data, produced estimates that did not differ materially from those reported in Table 3.

5. Discussion

The trial provides experimental evidence that the gamification design choice between individual leaderboards and cooperative team structures has consequences that depend on learner temperament. Introverted learners in the cooperative condition showed significantly higher engagement and persistence than their counterparts in the leaderboard condition, with moderate effect sizes, while extraverted learners were largely unaffected by the choice of design. The findings confirm the concern raised in our earlier study (Chen & Cheungsirakulvit, 2025) that leaderboard mechanics impose a temperament-related cost, and they demonstrate that this cost can be avoided by an alternative design without penalising other students.

The results are readily interpreted through self-determination theory (Deci & Ryan, 2000). Public ranking positions introverted learners in a continuous social comparison that they experience as threatening, undermining the sense of competence and autonomy that sustains effort; cooperative structures replace comparison with interdependence, satisfying relatedness and allowing competence to be experienced through contribution to a shared goal. The findings also align with the cooperative learning tradition (Johnson & Johnson, 2009; Slavin, 2014), which has long held that group goals with individual accountability promote effort, and with the field evidence of Morschheuser et al. (2019) that cooperative gamification generates more positive affect than competitive designs. The absence of a cost for extraverted learners is notable. It might have been expected that students who enjoy competition would lose motivation without a leaderboard, but their engagement and persistence were maintained, possibly because team-level goals still provided a target and because the social interaction inherent in teamwork is itself rewarding for extraverted students.

The achievement findings warrant careful interpretation. The overall effect in favour of the cooperative condition was small, and the interaction did not reach significance, so the trial does not establish that cooperative gamification improves learning specifically for introverted students. A ten-week unit may be too short for motivational differences to accumulate into achievement differences, and the end-of-unit test may be insensitive to the deeper engagement that cooperative structures appear to foster. Nevertheless, the direction of all effects was consistent, and there was no indication that cooperative gamification harmed achievement for any group, which addresses a common concern about cooperative structures in STEM subjects.

For practice, the findings suggest that leaderboards should be treated as an optional mechanic rather than a default in secondary STEM classrooms, and that teachers who adopt gamification should consider team-based structures as the baseline design. The cooperative condition in this trial was implemented with an existing platform and modest teacher training, indicating that the design is feasible at scale. Where leaderboards are retained, options such as private ranking, opt-in visibility or ranking relative to one's own prior performance may mitigate the social comparison effect (Festinger, 1954), although these variants were not tested here. More broadly, the study illustrates the value of examining individual differences in the evaluation of educational technology. Average effects reported in meta-analyses (Sailer & Homner, 2020) can conceal subgroups for whom a design is actively harmful, and Cain

(2012) is correct to note that such subgroups are often the quieter students whose disengagement is least visible.

6. Conclusion

This randomised controlled trial in 24 secondary STEM classrooms compared cooperative team-based gamification with individual leaderboards and found that introverted learners in the cooperative condition showed significantly higher engagement and persistence than introverted learners in the leaderboard condition, while extraverted learners performed comparably across conditions. Achievement effects were small and favoured the cooperative condition. The findings extend our earlier work on leaderboard-induced demotivation from diagnosis to remedy and support the recommendation that cooperative structures be adopted as the default gamification design in secondary STEM classrooms.

Limitations include the relatively short intervention period, the reliance on a single measure of achievement, the use of self-report engagement measures alongside log data and the confinement of the sample to one Canadian province. The trial also compared only two designs and cannot speak to hybrid mechanics. Future research should examine longer interventions, test modified leaderboard designs such as private or self-referenced ranking, replicate the trial in different cultural contexts where norms around competition and modesty differ, and investigate whether the engagement gains observed here persist beyond the gamified unit.

Acknowledgment

The author thanks the participating school districts, teachers and students, and acknowledges support from the University of British Columbia Faculty of Education.

References

Cain, S. (2012). *Quiet: The power of introverts in a world that can't stop talking.* Crown.

Chen, E. H., & Cheungsirakulvit, T. (2025). Gamification failure: Why leaderboards demotivate introverted learners in STEM classrooms. In *Proceedings of the 2025 International Conference on Creativity, Management, Education, Technology and Sciences* (pp. 212–220). Education Studio (Thailand). https://doi.org/10.62788/ICCMETS2025-paper21

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference* (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Eysenck, H. J. (1967). *The biological basis of personality.* Charles C. Thomas.

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

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

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

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Morschheuser, B., Hamari, J., & Maedche, A. (2019). Cooperation or competition – When do people contribute more? A field experiment on gamification of crowdsourcing. *International Journal of Human-Computer Studies, 127*, 7–24. https://doi.org/10.1016/j.ijhcs.2018.10.001

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.).* Sage.

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

Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicología, 30*(3), 785–791. https://doi.org/10.6018/analesps.30.3.201201