Effects of Teaching Mode on Physics Conceptual Mastery and Learning Engagement: A Comparison between Flipped and Traditional Classrooms

SONG JUNRAN[1] (Corresponding author, First author) SHI CHUCHU[1]
[1]School of Information Engineering, Jiangxi Science & Technology Normal University
589 Xuefu Avenue (Hongjiaozhou Campus), Nanchang 330038, Jiangxi, China.
E-mail: songjunran_0130@163.com

Abstract
This quantitative study examined whether teaching mode—flipped classroom (FC) versus traditional classroom (TC)—is associated with differences in high-school physics students perceived conceptual mastery (PCMS) and physics learning engagement (PLE). Using a cross-sectional, between-groups design, N = 298 students (FC n = 151; TC n = 147) completed brief Likert-type questionnaires (PCMS, 8 items; PLE, 8 items). Analyses were intentionally limited to descriptive statistics, independent-samples t tests, and one-way ANOVAs. Results showed that FC students reported significantly higher PCMS than TC students ($\Delta M = 0.22$, $t(296) = 3.21$, $p = .0015$, $d = 0.37$) and higher PLE ($\Delta M = 0.29$, $t(296) = 4.83$, $p < .001$, $d = 0.56$). Grade-level ANOVAs revealed modest developmental trends favoring upper grades (PCMS: $\eta^2 = .023$; PLE: $\eta^2 = .041$). Findings suggest that flipped organization is associated with moderately higher engagement and modestly higher perceived conceptual understanding beyond grade effects. Pedagogically, the results support dedicating in-class time to structured problem solving, guided inquiry, and formative feedback while moving content exposure to pre-class micro-materials. Limitations include the cross-sectional design with intact classes, reliance on self-report outcomes, and the absence of covariate adjustment. Future research should combine perceptual measures with standardized concept inventories (e.g., FCI/CSEM) and process data, and test design features and carefully scaffolded AI augmentations as potential moderators.
Keywords: flipped classroom; physics education; conceptual mastery; learning engagement; high school; independent-samples t test; one-way ANOVA
DOI: 10.7176/JEP/16-13-04
Publication date: December 30th 2025

1.Introduction

Over the past decade, the flipped classroom has matured from a niche instructional model into a widely adopted approach that redistributes lower-level cognitive processing (e.g., exposure to content) to pre-class time and reserves in-class time for higher-order learning activities. Syntheses in higher education show generally positive effects on learning outcomes and learner experience, while simultaneously flagging heterogeneity in design features, evaluation criteria, and contextual fit that complicate generalization across settings (Baig & Yadegaridehkordi, 2023; Oudbier et al., 2022). Reviews that span broader segments of the educational system further argue that flipped learning constitutes a transferable pedagogy yet emphasize persistent gaps in discipline-specific evidence outside university contexts (Galindo-Dominguez, 2021; Phillips & Wiesbauer, 2022).

Recent work has turned to technology-enhanced variants of flipped learning, demonstrating that AI-enabled personalization and data-driven supports can amplify engagement, motivation, and achievement when integrated with pre-class video and in-class active learning structures (Huang et al., 2023). In parallel, studies of prompt engineering for information retrieval suggest that large language models can streamline access to just-in-time resources within flipped workflows, potentially changing how learners prepare for and participate in class activities (Wang et al., 2023). While these developments outline an evolving ecosystem for flipped instruction, their consequences for secondary physics—a domain in which conceptual change and sustained engagement are both essential and notoriously difficult—remain under-specified.

Within physics education research (PER), decades of work have established concept inventories and simulation-based learning as effective levers for conceptual understanding, especially in mechanics and electricity and magnetism, with mounting evidence that interactive technologies can remediate misconceptions when thoughtfully embedded in pedagogy (Banda & Nzabahimana, 2021). Methodological advances such as eye-tracking have further illuminated how learners allocate attention and process visual-spatial information during physics tasks, reinforcing the importance of instructional designs that scaffold active cognitive processing rather than passive reception (Hahn & Klein, 2022). Against this backdrop, the flipped classroom appears theoretically well-aligned with the cognitive demands of physics, yet direct, controlled comparisons at the high-school level linking teaching mode to both conceptual mastery and learning engagement remain limited.

Concurrently, emergent studies have examined the roles of generative AI in physics learning contexts, reporting easy-to-implement classroom activities and cautioning that uncritical reliance on AI assistance can yield superficial problem solving and degraded conceptual reasoning (Bitzenbauer, 2023; Krupp et al., 2023). PER has also begun to explore data augmentation with large language models to support research workflows, underscoring both opportunities and validity concerns when AI interacts with physics content and assessment (Kieser et al., 2023). These findings sharpen a practical question for secondary physics teachers: does a flipped organization, even without sophisticated AI add-ons, measurably outperform traditional instruction on the dual targets of conceptual understanding and engagement?

The present study addresses this question by comparing flipped and traditional teaching modes in high-school physics using a cross-sectional, group-comparison design with questionnaire-based measures of conceptual mastery and learning engagement. Analyses are intentionally limited to descriptive statistics, independent-samples t tests, and one-way ANOVA, providing a transparent estimate of group differences that can inform subsequent, more elaborate causal designs. Specifically, we ask: (1) Do students in flipped classes report higher conceptual mastery than those in traditional classes? (2) Do students in flipped classes report higher learning engagement? (3) Are there subgroup differences by gender or grade that condition these effects? Consistent with prior syntheses on flipped learning and theory in PER, we hypothesize that flipped classes will show higher conceptual mastery and engagement than traditional classes, with possible moderation by student characteristics (Baig & Yadegaridehkordi, 2023; Galindo-Dominguez, 2021; Banda & Nzabahimana, 2021).

2.Methods

2.1Design and Participants

We employed a between-groups, cross-sectional comparative design contrasting flipped classrooms (FC) with traditional classrooms (TC) in high-school physics; intact classes from public schools offering comparable topics within the same 4–6-week window were cluster-sampled, with eligibility requiring parental consent/assent and attendance on survey day, and exclusions for >20% missing responses, straight-lining, or completion time <2 minutes; the target sample was N≈300 (≈150 per group) to ensure ≥.80 power for small-to-moderate effects (two-tailed α=.05), and subgroup descriptors (gender, grade) were collected for descriptive summaries and exploratory stratification only.

2.2Measures

Outcomes were (a) Perceived Physics Conceptual Mastery (PCMS; 8 items, two reverse-coded) and (b) Physics Learning Engagement (PLE; 8 items), each administered on 5-point Likert scales (1=strongly disagree…5=strongly agree) and scored as item means (higher=better), while teaching mode (FC vs. TC) was verified via teacher report and a 3-item fidelity checklist (pre-class materials, in-class active tasks, formative feedback); demographics included gender and grade, with optional descriptors (weekly study hours, prior physics grade band) used for descriptive statistics only.

2.3Procedure

After school recruitment and scheduling to align topical coverage across FC and TC classes, researchers briefed students on anonymity and voluntariness, obtained consent/assent, and administered a paper or secure online questionnaire during a single class session (~10 minutes) under standardized proctoring; completed forms were screened on-site for completeness and later programmatically checked for missingness patterns and aberrant response behavior.

2.4Data Preparation

Case-level exclusions followed pre-registered rules, item-level missingness ≤5% was imputed via person-mean within scale, scale-level missingness >20% was set to missing, outliers due to entry error were corrected or removed, and variables were coded as group (0=TC,1=FC), gender (0=male,1=female), grade (10/11/12), with outcome means PCMS_mean and PLE_mean computed after reverse-coding; all data were de-identified prior to analysis.

2.5Statistical Analysis

Analyses (SPSS 27) were limited to descriptive statistics (M, SD, 95% CI by group and total), independent-samples t tests comparing FC vs. TC on PCMS_mean and PLE_mean with Levene's test and Welch's t where appropriate, and one-way ANOVAs for grade-based differences (and, optionally, within-group ANOVAs by teaching mode), reporting F(df1,df2), p, η^2 (or ω^2), and Tukey or Games–Howell post-hoc tests as dictated by homogeneity; effect sizes (Cohen's d or Hedges' g) and mean differences with 95% CIs were reported for all primary contrasts.

2.6Ethics

Institutional approval was obtained prior to data collection; participation was voluntary with opt-out at any time, no personally identifying information was recorded, and results were reported in aggregate to protect confidentiality.

3.Findings

3.1 Descriptive Statistics

To profile the sample and provide an initial view of group differences, we report means and standard deviations for Perceived Physics Conceptual Mastery (PCMS) and Physics Learning Engagement (PLE). The analytic sample comprised N = 298 students, with flipped classrooms (FC) *n* = 151 and traditional classrooms (TC) *n* = 147; gender and grade distributions were comparable across groups.

Table 1. Descriptive statistics (M ± SD) for PCMS and PLE (three-line table)

Measure	Total (N=298)	FC (n=151)	TC (n=147)
PCMS (Conceptual Mastery)	3.67 ± 0.56	3.78 ± 0.52	3.56 ± 0.57
PLE (Learning Engagement)	3.77 ± 0.52	3.91 ± 0.49	3.62 ± 0.53

Note. Scale range = 1–5; higher scores indicate higher levels.

The FC group shows higher means than the TC group on both outcomes ($\Delta M \approx 0.22$–0.29), suggesting a potential teaching-mode advantage. Dispersion is moderate ($SD \approx 0.49$–0.57), indicating adequate variability for subsequent inferential tests.

Taken together, these descriptive patterns align with H1 (PCMS) and H2 (PLE) directionally; formal significance and effect sizes are evaluated with **t** tests and **ANOVA** in the following sections.

3.2 Between-Group Differences by Teaching Mode

We compared FC and TC on PCMS and PLE using independent-samples t tests. Levene's tests for homogeneity of variance were non-significant ($p > .05$), so equal-variance results are reported.

Table 2. Independent-samples t tests for teaching mode (FC vs. TC) (three-line table)

Measure	FC (M ± SD)	TC (M ± SD)	t(df)	p	Cohen's d	ΔM (95% CI)
PCMS	3.78 ± 0.52	3.56 ± 0.57	3.21 (296)	0.0015	0.37	0.22 [0.09, 0.36]
PLE	3.91 ± 0.49	3.62 ± 0.53	4.83 (296)	<0.001	0.56	0.29 [0.17, 0.41]

Note. ΔM = FC − TC; benchmarks for $d \approx 0.20/0.50/0.80 \approx$ small/medium/large.

Students in FC reported significantly higher conceptual mastery than those in TC ($d = 0.37$, small-to-medium), with the mean difference CI excluding zero. This suggests a robust advantage of flipped organization for perceived conceptual understanding.

For learning engagement, FC also outperformed TC with a medium effect ($d = 0.56$), indicating that flipped structures are associated with measurably higher behavioral/cognitive engagement in high-school physics. Both findings support H1 and H2.

3.3 Grade Differences

To examine developmental patterns, we ran one-way ANOVAs across grades (10/11/12) for PCMS and PLE in the full sample. Homogeneity tests were non-significant ($p > .05$); Tukey post-hoc tests were used.

Table 3. One-way ANOVA by grade (three-line table)

Measure	SS_between	SS_within	df1, df2	F	p	η²	Significant Tukey contrasts
PCMS	2.18	90.74	2, 295	3.55	0.030	0.023	Grade 12 > Grade 10 (ΔM = 0.18, p = 0.024)
PLE	3.41	79.92	2, 295	6.29	0.002	0.041	Grade 12 > Grade 10 (ΔM = 0.24, p = 0.003); Grade 11 > Grade 10 (ΔM = 0.15, p = 0.047)

Note. Rough benchmarks for $\eta^2 \approx .01/.06/.14 \approx$ small/medium/large.

ANOVA results indicate a significant grade effect on PCMS ($\eta^2 = .023$, small) with higher mastery in Grade 12 than Grade 10, consistent with cumulative exposure to physics curricula and practice.

For PLE, the grade effect is stronger ($\eta^2 = .041$, small-to-approaching-medium), with Grade 12 and Grade 11 both exceeding Grade 10. This pattern suggests that engagement tends to increase with seniority, potentially reflecting greater exam salience, strategy sophistication, or classroom norms.

While grade effects are statistically reliable, their magnitudes are modest, implying that classroom-level factors such as teaching mode remain key levers. Considering Section 4.2, the flipped advantage may add to natural grade-related gains; future work could test a two-factor design to probe potential interactions, though the present study intentionally reports only one-way ANOVAs per the preregistered analytic scope.

4.Discussion

This study compared flipped and traditional teaching modes in high-school physics using questionnaire indicators of perceived conceptual mastery (PCMS) and learning engagement (PLE), analyzed with descriptive statistics, independent-samples *t* tests, and one-way ANOVAs. The results showed statistically higher PCMS and PLE in flipped classrooms with small-to-medium effect sizes (PCMS: $d \approx 0.37$; PLE: $d \approx 0.56$), alongside modest grade trends favoring upper grades. Taken together, these findings align with meta-analytic and review evidence that flipped organization tends to improve learning outcomes and learner experience while leaving room for contextual variability (Baig & Yadegaridehkordi, 2023; Galindo-Dominguez, 2021; Oudbier et al., 2022; Phillips & Wiesbauer, 2022). The stronger effect on engagement than on perceived mastery is theoretically coherent: flipping reallocates class time toward active, collaborative, and problem-centered tasks, which are proximal drivers of behavioral/cognitive engagement and only indirectly translate to perceived conceptual gains.

From a physics education perspective, the pattern is consistent with research showing that interactive technologies and activity-rich pedagogies help surface and remediate misconceptions, particularly in mechanics and E&M, when embedded into structured tasks (Banda & Nzabahimana, 2021). The flipped format can be read as a delivery architecture that increases time-on-task for such cognitively demanding activities, while pre-class preparation reduces extraneous load during in-class sense-making. Our small-to-medium differences therefore appear educationally meaningful: even moderate gains in engagement often compound into improved study habits and strategic regulation over a term, which may precede measurable changes in conceptual inventories.

The grade effects (Grade 12 > Grade 10; Grade 11 > Grade 10 for PLE) likely reflect cumulative exposure to physics curricula, rising exam salience, and maturing study strategies. Importantly, the magnitudes ($\eta^2 \approx .02$–$.04$) indicate that teaching mode remains a tractable lever above and beyond developmental trends. Practically, the present evidence supports continuing or expanding flipped implementations in secondary physics, especially when class time can be guaranteed for structured problem-solving, guided inquiry, and formative feedback—design features repeatedly highlighted as critical to flipped efficacy (Baig & Yadegaridehkordi, 2023; Oudbier et al., 2022).

Technology choices surrounding flipped delivery deserve careful calibration. AI-enabled personalization has been shown to amplify engagement and motivation in flipped settings (Huang et al., 2023), and large-language-model–based prompt strategies can streamline pre-class resource access (Wang et al., 2023). At the same time, recent studies caution that uncritical reliance on generative AI can foster superficial solution patterns and undermine conceptual reasoning (Bitzenbauer, 2023; Krupp et al., 2023), and PER has begun to scrutinize the validity of AI-assisted data practices (Kieser et al., 2023). For secondary physics, the prudent stance is augmentation, not substitution: leverage AI to scaffold preparation and feedback while preserving cognitively productive struggle in class. Future classroom trials might also integrate process measures (e.g., classroom analytics or eye-tracking where feasible) to verify that flipped tasks indeed shift attention toward conceptually relevant cues (Hahn & Klein, 2022).

Several limitations temper causal inference. First, the design was cross-sectional and group-comparative with intact classes; although topical alignment and fidelity checks were implemented, unmeasured teacher/class factors may remain. Second, outcomes were self-report—appropriate for assessing perceived mastery and engagement but not interchangeable with performance; future studies should pair these with concept inventories (e.g., FCI/CSEM) and performance tasks to triangulate effects. Third, our analytic scope (descriptives, *t*, ANOVA) intentionally avoids covariate adjustment; subsequent work could use randomized or quasi-experimental pre–post designs, include fidelity/process covariates, and test moderation (e.g., gender, prior achievement) more formally. Finally, heterogeneous implementations of "flipped" models in the literature (Baig

& Yadegaridehkordi, 2023; Oudbier et al., 2022) underscore the need to report granular design features (video length, readiness checks, in-class task structure, feedback cycles).

In sum, the present evidence indicates that, in high-school physics, flipped classrooms are associated with higher learning engagement and modestly higher perceived conceptual mastery relative to traditional instruction. For practitioners, the immediate implication is to prioritize pre-class micro-materials and in-class, feedback-rich problem solving; for researchers, the next step is to combine perceptual outcomes with standardized concept measures and process data, and to probe design elements and AI augmentations that maximize engagement without eroding conceptual rigor.

5.Conclusion

This study compared flipped and traditional teaching modes in high-school physics using questionnaire measures of perceived conceptual mastery (PCMS) and physics learning engagement (PLE) and restricted analyses to descriptive statistics, independent-samples t tests, and one-way ANOVAs. Results showed that flipped classrooms were associated with moderately higher engagement and modestly higher perceived mastery than traditional classrooms, with small-to-medium effect sizes. Grade-level analyses indicated incremental gains in both outcomes from Grade 10 to Grade 12, yet the magnitudes suggested that teaching mode remains a meaningful, malleable lever beyond developmental trends.

Pedagogically, these findings support implementing flipped designs that shift lower-level content exposure to pre-class micro-materials and dedicate class time to structured problem-solving, guided inquiry, and formative feedback. Such designs are theoretically consonant with physics education principles that emphasize active sense-making and remediation of misconceptions. Practically, schools can prioritize short pre-class videos with readiness checks, in-class collaborative tasks targeting core concepts (e.g., mechanics, E&M), and rapid feedback cycles to convert time-on-task into sustained engagement.

Three constraints delimit interpretation. First, the cross-sectional, group-comparative design with intact classes does not establish causality. Second, outcomes were self-reported; triangulation with standardized concept inventories and performance tasks is needed. Third, our analytic scope excluded covariate adjustment and interaction modeling. Future work should employ pre–post or randomized/quasi-experimental designs, pair perceptual measures with concept inventories (e.g., FCI/CSEM) and process data (e.g., classroom analytics), and examine design features (video length, readiness checks, task structure) and carefully scaffolded AI augmentations as moderators. Despite these limits, the present evidence indicates that a well-implemented flipped organization is a promising, readily actionable pathway to enhance engagement and perceived conceptual understanding in secondary-school physics.

References

Baig, M. I., & Yadegaridehkordi, E. (2023). Flipped classroom in higher education: A systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education, 20*(1), 61.

Banda, H. J., & Nzabahimana, J. (2021). Effect of integrating physics education technology simulations on students' conceptual understanding in physics: A review of literature. *Physical Review Physics Education Research, 17*(2), 023108.

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology, 15*(3), ep430.

Galindo-Dominguez, H. (2021). Flipped classroom in the educational system. *Educational Technology & Society, 24*(3), 44–60.

Hahn, L., & Klein, P. (2022). Eye tracking in physics education research: A systematic literature review. *Physical Review Physics Education Research, 18*(1), 013102.

Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of artificial intelligence–enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education, 194*, 104684.

Kieser, F., Wulff, P., Kuhn, J., & Küchemann, S. (2023). Educational data augmentation in physics education research using ChatGPT. *Physical Review Physics Education Research, 19*(2), 020150.

Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K. E., Lukowicz, P., Kuhn, J., … Karolus, J. (2023). Unreflected acceptance—Investigating the negative consequences of ChatGPT-assisted problem solving in physics education. *arXiv*. arXiv:2309.03087.

Oudbier, J., Spaai, G., Timmermans, K., & Boerboom, T. (2022). Enhancing the effectiveness of flipped classroom in health science education: A state-of-the-art review. *BMC Medical Education, 22*(1), 34.

Phillips, J., & Wiesbauer, F. (2022). The flipped classroom in medical education: A new standard in teaching. *Trends in Anaesthesia and Critical Care, 42*, 4–8.

Wang, M., Wang, M., Xu, X., Yang, L., Cai, D., & Yin, M. (2023). Unleashing ChatGPT's power: A case study on optimizing information retrieval in flipped classrooms via prompt engineering. *IEEE Transactions on Learning Technologies, 17*, 629–641.