



RESEARCH ARTICLE

Integrating Communication and Cognitive Theories in Virtual Reality-Based Physics Learning: Achievement, Engagement, and Social Interaction in Senior High School

Putri Nur Aziza Sonia^{1*} and Wiwin Susanti²

^{1,2}SMK Pawiyatan Surabaya, Indonesia

*Corresponding author: putrinurazizasonia551@gmail.com.

Article History

Received: 17 October 2026

Revised: 21 October 2026

Accepted: 23 October 2026

Published: 26 October 2026

Publication Information

DOI: 10.36456/es7c2496

Volume / Issue: 3 / 1

Pages: 39–51

Keywords

Community of Inquiry, Cognitive Theory of Multimedia Learning, Physics Education, Quasi-Experimental Design, Student Engagement, Virtual Reality

ABSTRACT

This two-school quasi-experimental study examined a bundled immersive virtual reality (VR) instructional model for Grade 11 optics that combined structured communication with cognitive scaffolding. Of 120 students screened, 102 were allocated across four intact classes, and 60 complete cases (30 per condition) were analyzed after six 90-minute sessions. Achievement was assessed before and immediately after the intervention; engagement questionnaires and structured observations captured additional outcomes. Posttest achievement was higher in the integrated VR condition ($M = 82.63$, $SD = 6.91$) than in conventional instruction ($M = 74.28$, $SD = 7.44$), a mean difference of 8.35 points, $t(58) = 4.50$, $p < .001$, Cohen's $d = 1.16$. A baseline-adjusted analysis also favored the VR condition, $F(1, 57) = 20.45$, $p < .001$, partial eta squared = .264. Cognitive, affective, and behavioral engagement, as well as observed social-interaction quality, were also higher in the VR group. These results pertain to a combined instructional package rather than an isolated VR effect. The use of four intact classes, 50% complete-case retention, the short intervention, and the absence of delayed assessment constrain causal and longitudinal interpretation.

INTRODUCTION

Immersive virtual reality (VR) offers interactive three-dimensional experiences that are difficult to reproduce with textbooks or ordinary classroom demonstrations. In school settings, VR can support attention and situational interest, yet its effects on achievement depend on the design and duration of learning activities. Studies involving younger learners and reviews of school-based applications document promising engagement outcomes alongside mixed evidence for transfer and durable achievement (Boel et al., 2025; Chen et al., 2022; Hsu et al., 2025; Makransky et al., 2020; Maroukas et al., 2023). Thus, the educational question is not whether immersion is inherently superior but which instructional arrangements help students learn from it (Chen & Syu, 2024; Cheng et al., 2022).

High-school physics is an appropriate setting for this question. Geometric optics requires learners to coordinate rays, representations, angles, and abstract formulas; wave phenomena similarly require connecting dynamic patterns with symbolic explanations. Immersive manipulation may support conceptual visualization but can also increase task demands or distract students through novelty. Research on segmented VR activities, reflection, and learning aids indicates that prompts and guidance can help learners focus their cognitive processing (Huang et al., 2024; Klingenberg et al., 2023; Li et al., 2024; Wang et al., 2025).

A second design concern is communication. VR experiences do not automatically produce con-

structive peer exchange, explanation, or collaborative reasoning. Well-specified roles and facilitation may be needed to turn a shared experience into a learning opportunity (Alenezi, 2022; Han et al., 2022; Webb et al., 2022). The present study therefore considers a combined learning sequence that places immersive optics tasks within a deliberately organized cycle of teacher guidance, small-group dialogue, and reflective consolidation.

Theoretical Integration and Conceptual Framework

The Community of Inquiry (CoI) framework identifies social, teaching, and cognitive presence as interrelated dimensions of an educational experience (Garrison et al., 2000). In this study, communication presence is an instructional-design label for the coordinated use of social interaction, teaching facilitation, and sustained inquiry; it is not presented as a separately validated psychometric construct. Social presence informed peer discussion, teaching presence informed teacher prompts and feedback, and cognitive presence informed the explanation and reflection activities linking concrete experiences to scientific concepts.

The Cognitive Theory of Multimedia Learning (CTML) explains learning in terms of verbal and pictorial processing, limited capacity, and active selection, organization, and integration of information (Mayer, 2014). These premises motivated short immersive segments, guided questions, visual-to-symbolic explanations, and post-task summaries. In optics, such scaffolds can orient students to the relevant ray or wave features rather than leaving them to explore the simulated environment without a clear objective (Klingenberg et al., 2023; Wang et al., 2025).

Figure 1 presents the conceptual instructional design for the intervention. Its arrows represent proposed pedagogical relationships among immersive affordances, communication structures, cognitive scaffolds, and outcomes. Neither CoI presence dimensions nor cognitive load were measured as statistical mediators; Figure 1 therefore represents a design framework, not an empirically confirmed mediation model.

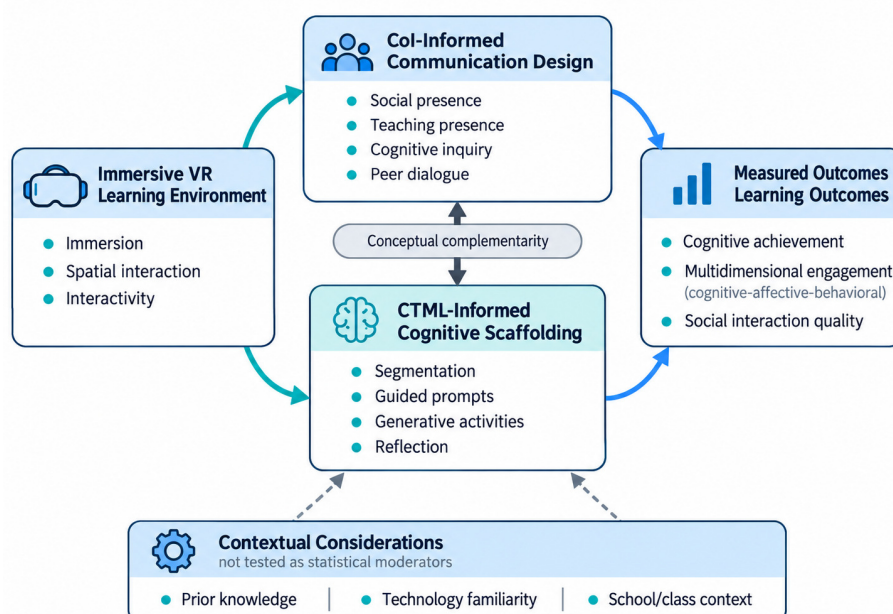


Figure 1. Conceptual communication-cognitive instructional design framework for immersive VR learning.

Note. Communication-presence structures and cognitive scaffolds were used to design the instructional condition. Illustrated pathways are theoretically proposed design relationships; no mediation paths or moderators were estimated.

Research Gap and Study Objectives

Previous VR research has examined immersion, cognitive learning aids, and collaboration, but relatively few school-based comparisons describe how these elements were combined and delivered in replicable classroom sequences (Hsu et al., 2025; Maroukias et al., 2023; Webb et al., 2022). This leaves a practical question about the outcomes associated with a theory-guided VR instructional package compared with a conventional sequence that pursues the same curriculum.

The study examined three immediate outcomes in Grade 11 physics: (1) post-intervention cognitive achievement, including short-term application tasks; (2) cognitive, affective, and behavioral engagement; and (3) observed social-interaction quality. It also examined engagement patterns across six sessions. The expectation was that outcomes would be more favorable in the bundled VR condition, although the design could not isolate headset immersion from the accompanying learning activities.

METHOD

Design, Setting, and Allocation

A two-arm, nonrandomized, quasi-experimental pretest–posttest design was implemented in two senior high schools during a two-week Grade 11 physics unit covering light propagation, reflection, refraction, interference, diffraction, and optical instruments. Allocation took place at the intact-class level, with two classes per school. One class in each school received the bundled VR model and the other received conventional instruction. Each school contributed 15 students to each condition, yielding 30 analyzed students per arm. Students were not individually randomized.

Two physics teachers participated, one in each school; each taught both instructional conditions at their school. The teachers completed eight hours of common lesson-plan and instructional-protocol preparation. This within-school arrangement reduced some between-school and teacher differences but did not remove class-level confounding. Because there were only four classes, outcomes analyzed at student level could not be treated as fully independent for strong causal inference.

Screening, Participant Flow, and Retention

Of 120 students screened, 18 were excluded before condition allocation: 11 because of self-reported severe motion sickness or vestibular susceptibility and seven because of uncorrected visual difficulties or problems fitting the headset. The remaining 102 students were allocated equally across the two instructional conditions ($n = 51$ each). The complete-case analysis included 60 students ($n = 30$ per group; ages 15–17 years). Figure 2 summarizes screening, condition allocation, exclusions, and the complete-case analysis.

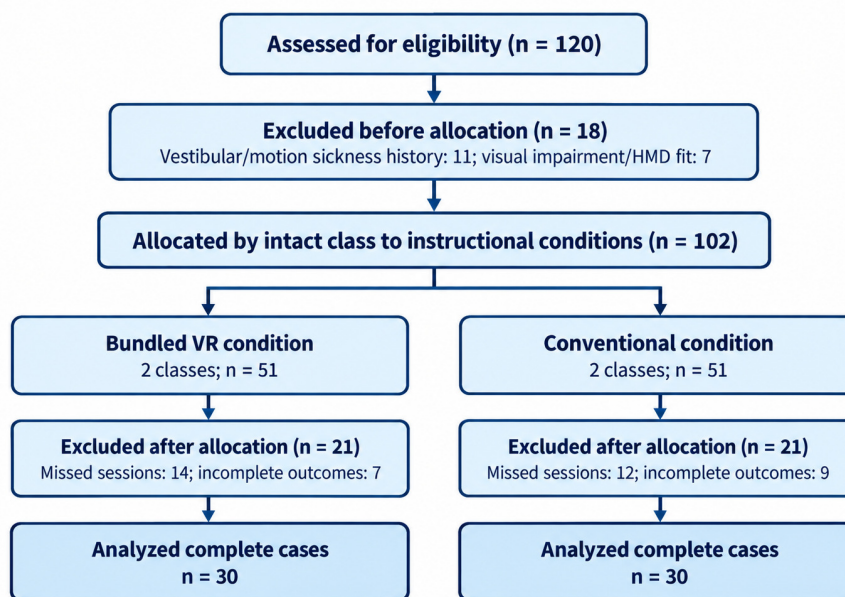


Figure 2. Participant screening, class allocation, attrition, and complete-case analysis.

Overall complete-case retention was 60/120 (50.0%) from screening and 60/102 (58.8%) after class allocation. The baseline-achievement comparison between included students ($M = 68.80$, $SD = 7.87$) and excluded students ($M = 67.95$, $SD = 8.12$) was not statistically significant, $t(118) = 0.58$, $p = .561$. This result indicates no statistically detectable mean difference in baseline achievement, but it does not establish that attrition was missing at random or rule out selection bias due to unmeasured characteristics.

Instructional Conditions and Equivalence

Both conditions addressed the same physics topics across six 90-minute sessions (540 minutes) and included small-group exercises. However, the instructional strategies were intentionally different: the experimental arm integrated immersive visualization, structured peer exchange, and multimedia-learning scaffolds, whereas the control arm relied on teacher explanation, textbook diagrams, worked problems, and conventional group discussion. The comparison therefore estimates differences associated with the instructional bundle, not a stand-alone VR hardware effect.

Table 1. Shared and condition-specific instructional characteristics

Parameter	Bundled VR instruction	Conventional instruction
Duration	6 × 90-minute sessions (540 min)	6 × 90-minute sessions (540 min)
Curriculum	Grade 11 optics: propagation, reflection, refraction, waves, instruments	Same learning objectives and topic order
Teacher contact	Direct briefing, guided prompts, and debriefing	Direct teaching, question answering, and review
Peer activity	Structured dialogue and 3–4-person problem solving	3–4-person textbook exercises and discussion
Feedback	Guided reflection and reciprocal peer feedback	Teacher explanations and peer answer checking
Digital immersion	Quest 2 interactive Unity simulation (up to 30 min)	No immersive headset exposure

Note. Matching applied to curricular topics and scheduled time; teacher–student activity and feedback mechanisms were not identical, by design.

Intervention Hardware, Sequence, and Fidelity

The immersive tasks were built in Unity 3D (version 2021.3 LTS) and presented on Oculus Quest 2 head-mounted displays (reported resolution 1,832 × 1,920 pixels per eye and 90 Hz refresh). An equipment orientation preceded the six instructional sessions. Each experimental session allocated 15 minutes for briefing, up to 30 minutes for the immersive task, 25 minutes for facilitated small-group exchange, and 20 minutes for guided reflection or consolidation; the VR exposure was capped at 30 minutes per session. Conventional lessons allocated 15 minutes for introductory teaching, 55 minutes for conventional instruction and exercises, and 20 minutes for discussion or synthesis.

Table 2. Six-session physics protocol for both instructional conditions

Session and physics focus	Learning target	Bundled VR activity (15/30/25/20 min)	Conventional sequence (15/55/20 min)
S1. Light propagation	Distinguish ray and wave representations	Ray manipulation, peer explanation, summary sheet	Ray-tracing diagrams, exercises, discussion
S2. Reflection and planar mirrors	Apply the law of reflection; describe plane-mirror images	Virtual-mirror lab and angle measurement, peer dialogue, reflection	Board demonstrations, protractor drawing, worksheet
S3. Refraction and lenses	Apply refraction principles; determine focal length and image formation	Virtual lens lab, refractive-index manipulation, guided problem solving	Lens formulas, textbook problems, group answer check
S4. Interference and diffraction	Explain phase and fringe spacing	Three-dimensional wave-pool task, peer negotiation, summarization	Two-dimensional diagrams, calculation, peer questions
S5. Optical instruments	Explain compound magnification	Multi-lens assembly, peer feedback cycle, reflective log	Static ray tracing, group synthesis
S6. Synthesis and review	Integrate optical models in novel problems	VR challenge lab, debriefing, synthesis activity	Worked practice examination, group review, summary

Note. Column headings express the sequence allocation in minutes: briefing / immersive or standard task / dialogue / consolidation for VR, and introduction / standard task / discussion for control. The instructional timing was kept consistent within each study condition.

An independent observer used an eight-item delivery checklist in all six sessions. The recorded fidelity percentages were 96.2% for the bundled VR sequence and 95.5% for conventional instruction. These percentages describe adherence to each condition's own protocol, rather than equivalence of instructional mechanisms. The checklist was intended to track adherence to the prescribed instructional sequence.

Measurement and Quality Evidence

Cognitive achievement was assessed using a 25-item optics test: 15 multiple-choice questions scored 0–1 and 10 short-answer application questions scored 0–2. The maximum raw score was 35, linearly transformed to a 0–100 scale. The blueprint assigned 15 items to knowledge/comprehension (C1–C2) and 10 to application and analysis (C3–C4). Three senior physics educators and two educational-technology experts reportedly evaluated its content alignment. Pretest and immediate posttest used the same assessment construct; there was no delayed retention assessment.

Kuder–Richardson Formula 20 (KR-20) coefficients were reported as .84 at pretest and .86 at posttest. Because the test includes 0–2 short-answer scores, the item-level coding behind KR-20 must be specified to establish whether these coefficients describe the full test or only dichotomously scored items. Reliability of the stated mixed-format total score was not separately reported.

Session engagement was assessed with a 15-item five-point questionnaire developed with reference to multidimensional engagement scholarship (Fredricks et al., 2004; Wang et al., 2016). Five items each addressed cognitive engagement (e.g., connecting optical ideas to real-life phenomena; $\alpha = .82$), affective engagement (interest and enjoyment; $\alpha = .86$), and behavioral engagement (participation in group problem solving; $\alpha = .79$); the reported full-scale coefficient was $\alpha = .88$. Questionnaires were completed after Sessions 1–6, and each student's session scores were averaged for the main between-group engagement comparisons. A separate general physics learning-orientation survey was reportedly administered before Session 1; its scores were not equivalent to the session-level outcome measure.

During instruction, social-interaction quality was rated on a 1–5 rubric covering peer-dialogue depth, reciprocal problem solving, and mutual feedback. The rubric drew on criteria for collaborative learning and on CoI dimensions (Garrison et al., 2000). Two research assistants reportedly completed four hours of calibration using sample classroom recordings. They were unaware of the specific study hypotheses, although the visible instructional condition prevented masking of head-set use. The reported ICC from a two-way random-effects model was .89 (95% CI [.82, .94]). Providing the rating anchors and rater-level protocol would strengthen replicability.

Table 3. Instruments, timing, and reported quality evidence

Measure	Structure or key indicators	Administration	Reported quality evidence
Achievement	15 × 1-point choice + 10 × 2-point short answer; 0–35 to 0–100	Pretest; immediate posttest	Five content reviewers; KR-20 .84/.86 (scoring basis unresolved)
Engagement	15 items; cognitive, affective, behavioral (5 each); 1–5	After each of six sessions	$\alpha = .82, .86, .79$; total $\alpha = .88$
Social interaction	1–5 observational rubric: dialogue, reciprocity, feedback	During sessions	Two raters; ICC = .89, 95% CI [.82, .94]
Baseline orientation	General engagement with physics prior to the unit	Before Session 1	Separate background measure; not a primary outcome

Statistical and Qualitative Analysis

Quantitative analyses were performed in SPSS version 26. Paired-samples t-tests described within-group achievement change, and independent-samples t-tests compared posttest achievement, aggregated engagement, and observed social interaction. Standardized posttest differences were expressed as Cohen's *d*, with 95% confidence intervals for principal between-group comparisons. A baseline-adjusted analysis of covariance (ANCOVA) compared immediate achievement with pretest achievement entered as a covariate; a two-group-by-six-session repeated-measures ANOVA examined overall engagement. The reported Shapiro–Wilk tests yielded *p* values above .10, and the Levene tests yielded *p* values above .15.

The analytic unit was the student although allocation was by intact class. With only two classes per condition and four classes overall, neither a defensible cluster-robust variance estimate nor a class-level random-effect analysis is supported by the available group summaries. The nominal individual-level *p* values and intervals are therefore descriptive or exploratory indicators and may be overly precise if within-class outcomes are correlated. Tests across multiple engagement domains were not adjusted for multiplicity. Sphericity diagnostics and any corrections to repeated-measures degrees of freedom were not reported.

Observation notes and open-ended student responses were reviewed following the six broad phases of thematic analysis described by Braun and Clarke (2006). Two coders achieved 86% raw agreement (Cohen's $\kappa = .79$). The resulting themes contextualized the quantitative outcomes but did not test causal mechanisms.

Ethics and VR Safety

The study received institutional ethics approval under protocol IEC-HS-2024-0412. Written informed consent was obtained from students and their guardians, and participants could withdraw without academic penalty. Research records were anonymized. Headset activities took place in a cleared area of approximately 2.5 × 2.5 m, with the device's Guardian boundary enabled. Participants were screened for susceptibility to motion sickness and visual discomfort, and VR exposure was capped at 30 minutes per session. Post-session symptoms were assessed using a modified Simulator Sickness Questionnaire. Three students experienced minor, transient eye fatigue during Session 1; the symptoms reportedly resolved after rest. No severe headset-related adverse events were reported.

RESULTS AND DISCUSSION

Participant Characteristics and Baseline Comparability

The final complete-case sample comprised 60 students, including 30 in each instructional condition and 15 per condition from each participating school. Mean pretest achievement was 68.47 (SD = 7.82) in the bundled VR condition and 69.12 (SD = 8.05) in the conventional condition. The baseline comparisons were not statistically significant ($p > .05$), although detailed group statistics for the general engagement-orientation survey were not available. Nonsignificant baseline differences do not establish individual or class equivalence; this study remains nonrandomized and vulnerable to class-level and selection effects.

Immediate Cognitive Achievement

Both groups improved on the immediate achievement test. The VR condition increased from 68.47 to 82.63 (mean gain 14.16), whereas the control condition increased from 69.12 to 74.28 (mean gain 5.16). The reported paired-tests were $t(29) = 9.84$, $p < .001$ for VR and $t(29) = 3.12$, $p = .004$ for conventional instruction (Table 4). These within-group tests describe change among complete cases and do not adjust for possible class-related dependence.

Table 4. Pretest and immediate posttest cognitive achievement

Condition	n	Pretest M (SD)	Posttest M (SD)	Mean gain	Paired t(29)	p
Bundled VR	30	68.47 (7.82)	82.63 (6.91)	14.16	9.84	< .001
Conventional	30	69.12 (8.05)	74.28 (7.44)	5.16	3.12	.004

Note. M = mean; SD = standard deviation. Gains and paired statistics describe change in the 60 complete cases, not an intention-to-treat sample.

The posttest mean difference of 8.35 points favored the integrated VR condition, $t(58) = 4.504$, $p < .001$, standardized $d = 1.163$ (Table 5). The pooled-variance t-statistic was consistent with the reported group means and standard deviations. The baseline-adjusted ANCOVA likewise favored the combined model, $F(1, 57) = 20.45$, $p < .001$, partial $\eta^2 = .264$; adjusted marginal posttest means were 82.72 (SE = 1.30) for VR and 74.19 (SE = 1.30) for conventional instruction. The reported homogeneity-of-slopes test was $F(1, 56) = 0.28$, $p = .598$. The ANCOVA is baseline-adjusted but does not remedy unmodeled class clustering.

Multidimensional Student Engagement

Averaging engagement responses across the six sessions produced higher mean scores in the bundled VR condition for cognitive engagement (4.21 vs 3.62), affective engagement (4.34 vs 3.71), behavioral engagement (4.18 vs 3.69), and overall engagement (4.24 vs 3.67). Table 5 reports the independent-samples comparisons across these domains. Reported standardized differences ranged from $d = 0.90$ for behavioral engagement to $d = 1.34$ for affective and overall engagement. These results describe students' responses during a highly structured learning package, not a separately estimated effect of immersion.

Table 5. Individual-level comparisons of achievement, engagement, and social interaction

Outcome	VR M (SD)	Control M (SD)	Difference [95% CI]	t(58)	p	d [95% CI]
Immediate achievement	82.63 (6.91)	74.28 (7.44)	8.35 [4.65, 12.05]	4.504	< .001	1.163 [0.61, 1.71]
Cognitive engagement	4.21 (0.48)	3.62 (0.55)	0.59 [0.32, 0.86]	4.427	< .001	1.143 [0.59, 1.69]
Affective engagement	4.34 (0.44)	3.71 (0.50)	0.63 [0.39, 0.87]	5.181	< .001	1.338 [0.77, 1.90]
Behavioral engagement	4.18 (0.52)	3.69 (0.57)	0.49 [0.21, 0.77]	3.479	.001	0.898 [0.36, 1.43]
Overall engagement	4.24 (0.39)	3.67 (0.46)	0.57 [0.35, 0.79]	5.177	< .001	1.337 [0.77, 1.90]
Social interaction	4.12 (0.46)	3.45 (0.51)	0.67 [0.42, 0.92]	5.343	< .001	1.380 [0.81, 1.94]

Note. CI = confidence interval; d = Cohen's d; posttest achievement is scored 0–100, other outcomes 1–5. Confidence intervals reflect student-level independent-samples comparisons. They do not account for class allocation or multiple testing, so precision may be overstated.

Descriptive overall engagement increased from $M = 3.98$ in Session 1 to $M = 4.31$ in Session

6 in the VR condition; control means were reported to lie between 3.60 and 3.74. The two-by-six repeated-measures analysis reported a group effect, $F(1, 58) = 26.80$, $p < .001$, partial $\eta^2 = .316$; a session effect, $F(5, 290) = 4.12$, $p = .001$, partial $\eta^2 = .066$; and a group-by-session interaction, $F(5, 290) = 3.85$, $p = .002$, partial $\eta^2 = .062$. The direction of the interaction was consistent with increasing reported engagement during the bundled VR sequence. The sphericity diagnostic and any corrected tests were not available, and students remained nested within intact classes.

Observed Social Interaction and Contextual Themes

The mean social-interaction rating was 4.12 (SD = 0.46) for bundled VR instruction and 3.45 (SD = 0.51) for conventional instruction. The individual-level comparison reported $t(58) = 5.343$, $p < .001$ and $d = 1.38$; the difference in means was 0.67 points (95% CI [0.42, 0.92]; Table 5). As shown in Figure 3, the scores differed by approximately two-thirds of a point on a five-point observation rubric. Since group condition was visually identifiable, expectancy and observer effects remain possible despite rater calibration.

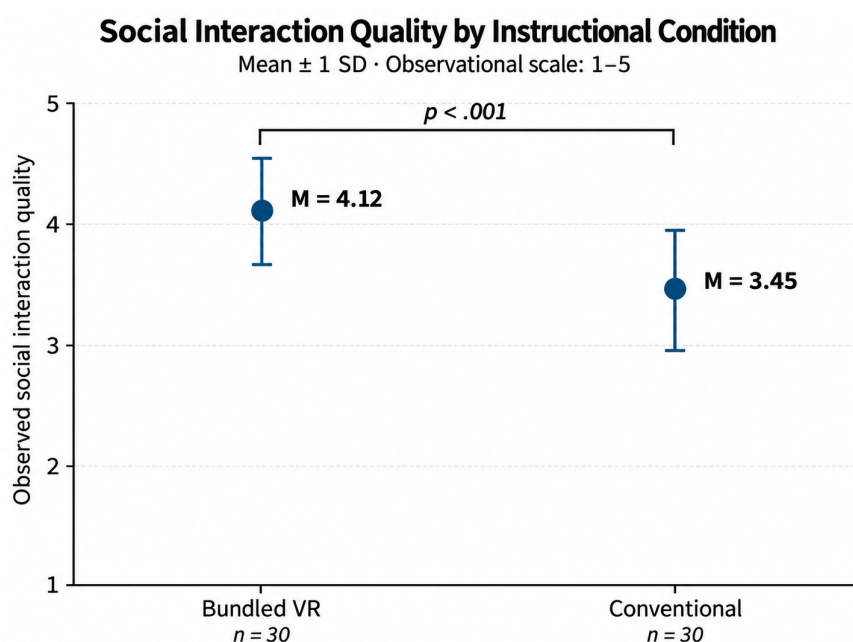


Figure 3. Mean observed social-interaction quality by instructional condition.

Note. Bars show condition means, and error bars indicate one standard deviation (1–5 scale). The comparison is based on complete cases and does not model class clustering.

Three themes from the reported qualitative review added descriptive context: (1) shared spatial representations appeared to make students' mutual explanations more specific; (2) guided prompts and reflection sheets helped students relate immersive tasks to abstract physics relationships; and (3) interest in the immersive simulations appeared to support persistence during challenging calculations. The themes are descriptive and should not be interpreted as independently tested pathways linking instructional design to achievement.

Interpretation of the Bundled Model

The pattern across immediate achievement, reported engagement, and observed social interaction is consistent with the proposition that visualization, collaborative explanation, and reflective

consolidation can work together in short physics learning sequences. Under CTML, segmentation and prompted explanations could help students organize dynamic optical representations (Klingenberg et al., 2023; Mayer, 2014). Under CoI, facilitation and reciprocal discussion could make learning activity more dialogic and purposeful (Garrison et al., 2000; Han et al., 2022; Webb et al., 2022). However, those explanations are theoretically plausible rather than experimentally tested mechanisms because neither cognitive load nor CoI presence was measured with dedicated mediator instruments.

The control group also showed meaningful gains, underscoring that conventional instruction remained productive. The design intentionally changed several elements together: medium, activity sequencing, prompts, feedback, and peer interaction. It is therefore inappropriate to attribute the numerical advantage specifically to VR hardware or immersion. This distinction matters because other studies report elevated presence or enthusiasm without consistent achievement transfer, as well as contextual variation across schools and affectively intensive learning tasks (Kaplan-Rakowski & Gruber, 2024; Makransky et al., 2020; Nachtigall et al., 2024; Šikl et al., 2024).

The engagement results may also reflect initial excitement about unfamiliar headsets. Although orientation reduced operating demands and session engagement rose during the intervention, a six-session series cannot distinguish a sustained pedagogical benefit from novelty, familiarity, or other time-varying influences. Studies of technology familiarization and immersive novelty similarly caution against making long-term motivational claims from short exposures (Miguel-Alonso et al., 2024; Rodrigues et al., 2022). No evidence of durable retention, long-term transfer, or post-intervention trajectories was collected; evidence from procedural VR transfer tasks is not directly generalizable to school optics learning (Petersen et al., 2022).

Practical and Theoretical Implications

For Grade 11 optics instruction, the design offers a potentially replicable sequence: state the objective, cap headset exposure, pose focused interactive tasks, require small-group explanations, and end with a written consolidation product. The sequence may be useful when students need to connect three-dimensional phenomena with ray diagrams and formulas. Teachers require competence in prompt design, feedback, and classroom organization as well as headset operation; implementation feasibility also depends on device cost, technical support, and learner safety (Schnyder et al., 2025).

Conceptually, the contribution is an explicit integration of communication-presence design and CTML-informed cognitive scaffolding. Figure 1 articulates an intervention logic that future research could test, but the current evidence does not validate a new construct or demonstrate mediation. More informative comparisons would separate immersive medium, communication structures, and cognitive supports in a multi-arm or factorial design, preferably with class-level randomization and a larger number of clusters.

Limitations and Future Research

Several limitations substantially constrain inference. First, only 60 of 120 screened students entered the complete-case analysis, and loss of participants at multiple stages may affect sample representativeness. A nonsignificant baseline comparison cannot prove missing-at-random attrition or eliminate selection bias. Second, treatment assignment was made at the intact-class level in only four classes from two schools. Even with within-school matching of teacher and subject mat-

ter, class dynamics and other unobserved factors can confound treatment estimates, while student-level t-tests, confidence intervals, repeated-measures ANOVA, and ANCOVA omit within-class dependence.

Third, the intervention was a bundled theory-guided package; no component effect or mediation estimate can be identified. Fourth, several instrument details require further documentation, particularly mixed-format test reliability, full scoring rubrics, raw observer agreement data, and baseline engagement scores. Fifth, the six-session intervention included no delayed assessment, and repeated engagement measurements may be affected by novelty, response bias, or unverified sphericity assumptions. Finally, the available analyses did not control the familywise error rate across outcomes or document a replication dataset.

Future investigations should pre-register class allocation, increase the number of schools and clusters, report participant-level attrition by cause and condition, and analyze data using methods that account for allocation. Multi-arm comparisons of VR with and without communication/cognitive supports, delayed transfer testing, objective discourse measures, complete psychometric reporting, and transparent de-identified datasets would clarify which mechanisms and benefits are durable.

CONCLUSION

In the studied two-school Grade 11 physics setting, a six-session instructional package combining immersive optics simulations, structured communication, and cognitive scaffolding was associated with higher immediate achievement, engagement, and observed social-interaction scores than conventional instruction. The reported baseline-adjusted achievement comparison was compatible with the unadjusted difference, but these student-level analyses cannot rule out bias from intact-class allocation, selective retention, or differences in the broader instructional package.

The study supports a cautious pedagogical proposition: immersive tools should be evaluated as part of a designed learning sequence, not assumed effective merely because they provide presence or interactivity. The findings do not establish isolated VR superiority, confirmed theoretical mediation, or long-term retention. Larger cluster-aware studies with transparent instrument validation and delayed follow-up are needed before those broader conclusions can be drawn.

DECLARATIONS

Funding

No external funding was reported for this study.

Conflict of Interest

The authors declare no conflict of interest related to this study.

Informed Consent

Written informed consent was obtained from participating students and their guardians. Participation was voluntary, and students could withdraw without academic penalty.

Data Availability Statement

De-identified participant-level data may be requested from the corresponding author for legitimate research purposes, subject to institutional approval and participant confidentiality. The dataset and SPSS syntax are not linked to a public repository in this article.

Declaration of AI Use

The authors used ChatGPT 5.6 to assist with language editing, manuscript organization, and document formatting. The authors remain responsible for verifying the study records, statistical outputs, ethics documentation, references, and the accuracy and integrity of the submitted manuscript.

REFERENCES

- Alenezi, A. R. (2022). Modeling the social factors affecting students' satisfaction with online learning: A structural equation modeling approach. *Education Research International*, 2022(1), 2594221. <https://doi.org/10.1155/2022/2594221>
- Boel, C., Rotsaert, T., Valcke, M., & Schellens, T. (2025). A sequencing effect of adding generative learning strategies to immersive virtual reality on middle school students' performance, intrinsic motivation, and self-efficacy. *Journal of Computer Assisted Learning*, 41(3), e70045. <https://doi.org/10.1111/jcal.70045>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chen, C.-H., & Syu, J.-Y. (2024). Effects of integrating a role-playing game into a virtual reality-based learning approach on students' perceptions of immersion, self-efficacy, learning motivation and achievements. *British Journal of Educational Technology*, 55(5), 2339–2356. <https://doi.org/10.1111/bjet.13436>
- Chen, Y.-T., Li, M., Huang, C.-Q., Han, Z.-M., Hwang, G.-J., & Yang, G. (2022). Promoting deep writing with immersive technologies: An SVVR-supported Chinese composition writing approach for primary schools. *British Journal of Educational Technology*, 53(6), 2071–2091. <https://doi.org/10.1111/bjet.13247>
- Cheng, K., Lee, S. W., & Hsu, Y.-T. (2022). The roles of epistemic curiosity and situational interest in students' attitudinal learning in immersive virtual reality environments. *Journal of Educational Computing Research*, 61(2), 494–519. <https://doi.org/10.1177/07356331221121284>
- 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>
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2–3), 87–105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Han, E., Nowak, K. L., & Bailenson, J. N. (2022). Prerequisites for learning in networked immersive virtual reality. *Technology, Mind, and Behavior*, 3(4), 566–579. <https://doi.org/10.1037/tmb0000094>
- Hsu, T.-C., Wen, W.-N., Liao, C.-S., Tu, Y.-F., & Lee, M.-J. (2025). Virtual reality in P-12 education for improving presence, immersion, and 4C skills: A systematic review of empirical research. *Thinking Skills and Creativity*, 58, 101918. <https://doi.org/10.1016/j.tsc.2025.101918>
- Huang, Y.-M., Wang, W.-S., Lee, H.-Y., Lin, C.-J., & Wu, T.-T. (2024). Empowering virtual reality with feedback and reflection in hands-on learning: Effect of learning engagement and higher-order thinking. *Journal of Computer Assisted Learning*, 40(4), 1413–1427. <https://doi.org/10.1111/jcal.12959>
- Kaplan-Rakowski, R., & Gruber, A. (2024). An experimental study on reading in high-immersion virtual reality. *British Journal of Educational Technology*, 55(2), 541–559. <https://doi.org/10.1111/bjet.13392>
- Klingenberg, S., Fischer, R., Zettler, I., & Makransky, G. (2023). Facilitating learning in immersive virtual reality: Segmentation, summarizing, both or none? *Journal of Computer Assisted Learning*, 39(1), 218–230. <https://doi.org/10.1111/jcal.12741>
- Li, Y.-F., Guan, J.-Q., Wang, X.-F., Chen, Q., & Hwang, G.-J. (2024). Examining students' self-regulated learning processes and performance in an immersive virtual environment. *Journal of Computer Assisted Learning*, 40(6), 2948–2963. <https://doi.org/10.1111/jcal.13047>
- Makransky, G., Petersen, G. B., & Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology*, 51(6), 2079–2097. <https://doi.org/10.1111/bjet.12954>
- Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). Virtual reality in education: A review of learning theories, approaches and methodologies for the last decade. *Electronics*, 12(13), 2832. <https://doi.org/10.3390/electronics12132832>

- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Miguel-Alonso, I., Checa, D., Guillen-Sanz, H., & Bustillo, A. (2024). Evaluation of the novelty effect in immersive virtual reality learning experiences. *Virtual Reality*, 28, 27. <https://doi.org/10.1007/s10055-023-00926-5>
- Nachtigall, V., Yek, S., & Rummel, N. (2024). Improving students' processing of history-related 360° videos through collaboration and emotion regulation. *Journal of Computer Assisted Learning*, 40(6), 3338–3359. <https://doi.org/10.1111/jcal.13076>
- Petersen, G. B., Klingenberg, S., & Makransky, G. (2022). Pipetting in virtual reality can predict real-life pipetting performance. *Technology, Mind, and Behavior*, 3(3), 452–467. <https://doi.org/10.1037/tmb0000076>
- Rodrigues, L., Pereira, F. D., Toda, A. M., Palomino, P. T., Pessoa, M., Carvalho, L. S. G., Fernandes, D., Oliveira, E. H. T., Cristea, A. I., & Isotani, S. (2022). Gamification suffers from the novelty effect but benefits from the familiarization effect: Findings from a longitudinal study. *International Journal of Educational Technology in Higher Education*, 19, 13. <https://doi.org/10.1186/s41239-021-00314-6>
- Schnyder, S., Dubach, J., Dall'Olio, L., Tempelmann, S., Cacchione, T., & Martarelli, C. S. (2025). Are primary schools ready for immersive virtual reality? Resistance among stakeholders. *Humanities and Social Sciences Communications*, 12, 1304. <https://doi.org/10.1057/s41599-025-05702-1>
- Šikl, R., Brücknerová, K., Švedová, H., Děchtěrenko, F., Ugwitz, P., Chmelík, J., Pokorná, H., & Juřík, V. (2024). Who benefits and who doesn't in virtual reality learning: An experimental study comparing two types of school. *Journal of Computer Assisted Learning*, 40(4), 1591–1604. <https://doi.org/10.1111/jcal.12973>
- Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, 43, 16–26. <https://doi.org/10.1016/j.learninstruc.2016.01.008>
- Wang, W.-S., Lee, H.-Y., Lin, C.-J., Li, P.-H., Huang, Y.-M., & Wu, T.-T. (2025). Enhancing students' learning outcomes in self-regulated virtual reality learning environment with learning aid mechanisms. *British Journal of Educational Technology*, 56(1), 366–387. <https://doi.org/10.1111/bjet.13512>
- Webb, M., Tracey, M., Harwin, W., Tokatli, O., Hwang, F., Johnson, R., Barrett, N., & Jones, C. (2022). Haptic-enabled collaborative learning in virtual reality for schools. *Education and Information Technologies*, 27(1), 937–960. <https://doi.org/10.1007/s10639-021-10639-4>