

Enhancing Learning Outcomes and Engineering Self-Efficacy in Digital Logic Design through Embedded Formative Assessment: A Quasi-Experimental Study

Peihe Liu, Zixuan Wang*, Jianxin Wang, Runze Zhou, Chao'en Xiao, Lei Zhang

Beijing Electronic Science and Technology Institute, China

**Corresponding author:* Zixuan Wang, 20253811@mail.besti.edu.cn

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Traditional digital logic design courses rely heavily on lecture-dominated instruction with terminal examinations, providing little opportunity for timely formative feedback or the development of learner self-efficacy. This study investigated whether a blended instructional model incorporating embedded formative assessment—delivered through weekly low-stakes online quizzes, structured pre-lab prediction tasks, and instructor-facilitated peer critique sessions—could improve learning outcomes and self-efficacy relative to conventional instruction. A quasi-experimental study was conducted across three cohorts ($n=176$) over one academic year at a research-intensive university in China. Outcome measures included a validated Digital Logic Achievement Test (DLAT), the Domain-Specific Self-Efficacy scale (DSSE), and final course grades. The blended cohort significantly outperformed controls on the DLAT post-test (Cohen's $d=0.54$) and on the DSSE ($d=0.67$), with a substantially lower failure rate (4.8% vs. 13.2%). Thematic analysis of post-course interviews ($n=18$) identified timely corrective feedback and progressive task difficulty as the principal mechanisms supporting both knowledge consolidation and self-efficacy growth. These findings suggest that moderate structural modifications to assessment frequency and delivery can meaningfully improve conceptual learning and motivational engagement in digital logic courses.

Keywords: Blended instruction; Retrieval practice; Peer critique; Classroom feedback; Undergraduate STEM education; Student motivation

Online publication: September 30, 2026

1. Introduction

Digital Logic Design (also known as Digital Logic Fundamentals or Digital Circuits) is a compulsory disciplinary foundation course for students in electronic information engineering, communication engineering, and computer science programs. Its content spans Boolean algebra, combinational circuit

analysis and design, sequential logic, finite-state machines (FSMs), and programmable logic devices—a knowledge system distinguished by high theoretical abstraction and strong practical integration^[1]. Despite its structural importance as a gateway to advanced courses such as microprocessor principles and embedded systems design, its abstract, hierarchically organized content has been associated with elevated attrition in undergraduate STEM programs^[2] and imposes a high cognitive load on novice learners^[3].

The persistence of poor learning outcomes in digital logic courses is not attributable solely to content difficulty. Substantial evidence points to structural features of conventional instruction as a proximate cause: specifically, the reliance on large-group lectures punctuated by infrequent, high-stakes assessments that return feedback too slowly to guide ongoing learning^[4]. Under this model, students typically discover conceptual misunderstandings only at or after the final examination—far too late for meaningful corrective action. Cognitive science provides a clear explanation: durable understanding requires repeated retrieval, elaboration, and error correction distributed across the learning period, not massed immediately before a terminal test^[5,6].

Formative assessment—assessment designed explicitly to inform and adjust instruction or learning during the instructional process—is one of the most robustly supported tools for achieving these distributed practice conditions^[7,8]. When formative assessments are designed with low stakes, immediate feedback, and progressive difficulty, they not only consolidate procedural and conceptual knowledge but also build learner self-efficacy: the domain-specific belief in one’s capacity to perform required tasks^[9,10]. Self-efficacy, in turn, is a powerful predictor of persistence, strategy use, and ultimate achievement in technically demanding STEM courses^[11,12].

Blended instruction—the purposeful integration of online and face-to-face components—provides a practical vehicle for embedding formative assessment activities without prohibitively increasing instructor workload^[13]. Online quiz platforms enable automated scoring and immediate item-level feedback; structured pre-lab prediction tasks make tacit mental models explicit before students encounter hardware; and peer critique sessions generate explanatory dialogue that consolidates understanding for both reviewer and reviewed^[14,15]. Yet despite strong evidence that active-learning approaches outperform conventional instruction across engineering disciplines^[16], the systematic use of embedded formative assessment within digital logic courses specifically remains limited.

This study therefore addressed three research questions. RQ1: Does a blended instructional model with embedded formative assessment improve short-term learning achievement relative to conventional instruction? RQ2: Does it improve domain-specific self-efficacy? RQ3: Do achievement gains persist over an eight-week retention interval? Through a quasi-experimental design spanning three cohorts ($n=176$), we provide quantitative and qualitative evidence bearing on these questions and derive practical design principles for instructors seeking evidence-based alternatives to terminal-examination-dominated digital logic instruction.

2. Theoretical framework and related work

2.1. Formative assessment in engineering education

Formative assessment has been extensively studied in K-12 and general higher-education settings, where meta-analyses of feedback and formative practices consistently report medium positive effects on achievement (typically around $d=0.5$)^[7,8]. Evidence from engineering education is more recent but converging:

active-learning and peer-instruction methods have been associated with substantial gains in examination performance across STEM disciplines ^[16,17], and structured formative feedback has been linked to deeper conceptual revision ^[18]. Within engineering education specifically, prior reform efforts have concentrated on curriculum-level competency mapping ^[19] and laboratory restructuring such as inductive, project-based FPGA work ^[20], while the role of routine within-semester formative assessment frequency in digital logic courses has received comparatively little systematic attention.

2.2. Self-efficacy in STEM learning

Bandura's social cognitive theory holds that self-efficacy—an individual's judgment of capability to organize and execute courses of action required to produce given outcomes—is shaped principally by mastery experiences, vicarious observation, social persuasion, and physiological states ^[9]. In STEM contexts, domain-specific self-efficacy predicts academic persistence, choice of learning strategies, and performance independent of prior achievement ^[11,12]. Formative assessment creates mastery experiences: when a student successfully retrieves a previously confused concept, the experience incrementally raises self-efficacy for that domain. Peer critique sessions provide vicarious learning: observing a peer explain a circuit behavior raises the observer's belief that the task is achievable ^[10]. A blended design that systematically incorporates both mechanisms is therefore theoretically positioned to produce self-efficacy gains beyond what either element would yield in isolation.

2.3. Blended learning in undergraduate STEM

A recent meta-analysis by Yu *et al.* found that blended learning produced an upper-medium positive effect on student learning performance (Hedges' $g=0.65$) relative to traditional instruction, with the magnitude of the effect moderated by the teaching method and the proportion of the course delivered online ^[27]; this is consistent with broader arguments for the transformative potential of well-designed blended environments ^[13]. The present study implements this principle by ensuring that each of the three embedded formative assessment types targets a distinct cognitive process: retrieval practice (quizzes), anticipatory prediction error correction (pre-lab tasks), and elaborative explanation (peer critique), and that all three are explicitly coordinated with weekly lecture content rather than added as parallel, disconnected activities.

3. Research design and methods

3.1. Study context and participants

The study was conducted in the Digital Logic Design course at a research-intensive university in central China. The course runs for 64 hours per semester (48 lecture + 16 laboratory) and is compulsory for second-year students in electronic information engineering and communication engineering. Three cohorts were recruited across two consecutive semesters of the 2023–2024 academic year. Two cohorts from Semester 1 ($n=87$) served as the control condition, receiving conventional lecture-and-laboratory instruction with biweekly homework assignments and one mid-term and one final examination. One cohort from Semester 2 ($n=89$) received the blended-FA intervention. Baseline characteristics are presented in **Table 1**.

To minimize confounding, the two conditions were held identical on every dimension except the embedded formative assessment activities. The instructor was identical—the same lead instructor delivered all lectures and laboratory sessions across both conditions, eliminating teacher-effect confounds. The

materials were identical—both cohorts used the same textbook, lecture slides, laboratory manuals, and problem sets. The summative assessment was identical—both conditions sat the same mid-term and final examinations, marked against the same rubric by the same graders, who were blind to condition. The schedule was identical—both conditions followed the same 16-week timetable, the same topic sequence, the same 48 lecture hours, and the same 16 laboratory hours. The sole systematic difference between conditions was the presence of the three formative assessment components in the blended-FA cohort.

Table 1. Baseline characteristics of the control and blended-FA cohorts

Characteristic	Control (<i>n</i> = 87)	Blended-FA (<i>n</i> = 89)	<i>P</i> -value
Age, mean ± SD (years)	19.5 ± 0.7	19.4 ± 0.6	0.41
Prior GPA (out of 4.0)	2.87 ± 0.53	2.90 ± 0.49	0.68
DLAT pre-test, mean ± SD (%)	50.8 ± 12.1	51.2 ± 11.8	0.82

Independent-samples *t*-tests confirmed no significant between-group differences on any baseline variable (all *P* > 0.05), supporting the assumption of group equivalence prior to the intervention.

3.2. Intervention description

The blended-FA intervention embedded three interleaved formative assessment components within the existing course structure without altering contact hours or the syllabus. The components and their feedback mechanisms are summarized in **Table 2**, and the overarching conceptual model is illustrated in **Figure 1**.

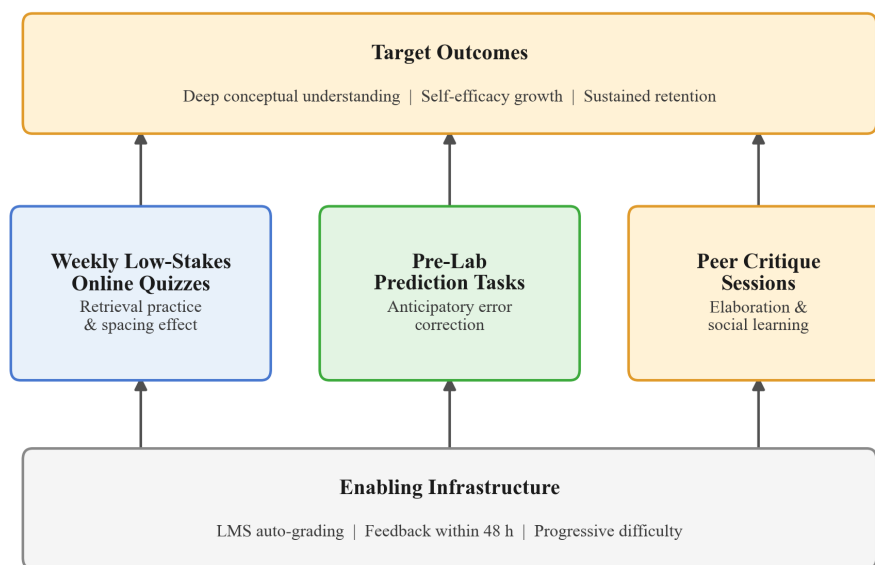


Figure 1. Conceptual model of the blended-FA intervention. The three embedded formative assessment components—weekly online quizzes, pre-lab prediction tasks, and peer critique sessions—converge on the targeted outcomes of deep conceptual understanding, high domain-specific self-efficacy, and sustained retention, supported by a common enabling infrastructure of timely feedback and progressive task difficulty.

Table 2. Components of the embedded formative assessment intervention

Component	Frequency and format	Feedback mechanism
Weekly online quizzes (10 items, 15 min)	Once per week via the course LMS, administered on Sunday evenings	Automated item-level feedback with worked explanations; visible within 2 min of submission
Pre-lab prediction tasks (1 short-answer item)	Before each of 8 laboratory sessions; submitted online 24 h in advance	Instructor written feedback within 48 h; discussed collectively in the first 5 min of lab
Peer critique sessions (structured protocol, 60 min)	Three times per semester; dyads exchange design solutions and complete a structured critique form	Written peer comments returned within one week; instructor-moderated whole-class debrief

Quiz difficulty was calibrated to the preceding week's lecture content and scaled progressively across the semester following Bloom's taxonomy: Levels 1–2 (recall and comprehension) in weeks 1–4, Levels 3–4 (application and analysis) in weeks 5–10, and Levels 5–6 (synthesis and evaluation) in weeks 11–14. Quizzes contributed 10% of the final grade in aggregate, keeping stakes low enough to encourage honest engagement and reduce test anxiety ^[6]. Pre-lab prediction tasks required students to predict the output waveform or state sequence of a specified circuit before constructing or simulating it, activating anticipatory knowledge and priming attention to disconfirming evidence ^[5]. Peer critique sessions used a structured protocol in which each student completed a standardized form covering correctness, clarity of design rationale, and suggestions for improvement, then participated in a 15-minute paired discussion.

3.3. Measurement instruments

The Digital Logic Achievement Test (DLAT) is a 40-item researcher-developed instrument covering five content domains: number systems and Boolean algebra (8 items), combinational logic analysis and design (10 items), sequential circuit design (12 items), finite-state machines (6 items), and programmable logic devices (4 items). It was constructed following established principles of criterion-referenced achievement-test design ^[21] and piloted with a separate cohort the preceding semester to establish reliability (Cronbach's $\alpha=0.84$; two-week test-retest reliability=0.79). The DLAT was administered as a pre-test (Week 2), post-test (Week 17), and an eight-week retention test using a 20-item parallel-form subset matched for domain coverage and difficulty.

The Domain-Specific Self-Efficacy scale (DSSE) is a 12-item Likert instrument (1 = not at all confident; 5 = completely confident) adapted from Pajares and Miller's mathematics self-efficacy scale for the digital logic domain ^[22]. Sample items include: "I am confident that I can independently design a 4-bit synchronous counter using D flip-flops" and "I can identify the cause of a race condition in a combinational circuit without assistance." Internal consistency in a pilot study was $\alpha=0.87$. The adapted items were reviewed by three engineering education experts to ensure content validity for the digital logic domain. The DSSE was administered at the start and end of the course.

3.4. Data collection and analysis

Quantitative data were analyzed using SPSS 26. Between-group differences on continuous outcome measures were first examined with independent-samples *t*-tests, with effect sizes reported as Cohen's *d*. To adjust statistically for any residual baseline variation and to serve as the primary confirmatory test of the intervention effect, a one-way analysis of covariance (ANCOVA) was conducted on each post-test outcome, entering condition as the fixed factor and the corresponding pre-test score (DLAT or DSSE) as the covariate. The homogeneity-of-regression-slopes assumption was verified via the condition $\times$ covariate interaction

and held in every model (all $P > 0.20$); partial eta-squared (partial η^2) is reported as the ANCOVA effect-size index. Grade distribution differences were assessed with chi-square tests. Levene's test verified homogeneity of variance prior to each t -test, with Welch's correction applied where violated. All tests were two-tailed at $\alpha = 0.05$.

Eighteen students (9 per condition; stratified by achievement tercile) participated in individual semi-structured post-course interviews (25–35 min each), audio-recorded and transcribed verbatim. Transcripts were analyzed using reflexive thematic analysis ^[23]. Two researchers coded independently; intercoder agreement was $\kappa = 0.78$, with disagreements resolved by discussion.

3.5. Ethics statement

This study was conducted within routine classroom instruction and involved only the analysis of anonymized course data and voluntary questionnaires, posing minimal risk to participants. All participants were informed of the purpose of the study and provided informed consent; participation was voluntary and non-participation did not affect course grades. Students in the control condition were offered access to the blended-FA materials after data collection was complete, and all interview data were de-identified before analysis.

4. Results

4.1. Learning achievement (RQ1 and RQ3)

Table 3 presents descriptive statistics and effect sizes for the DLAT across all three measurement points. **Figure 2(a)** visualizes the score distributions.

Table 3. DLAT scores and retention results by condition (mean $\pm$ SD)

Measure	Control ($n=87$)	Blended-FA ($n=89$)	t (174)	Cohen's d
Pre-test (%)	50.8 $\pm$ 12.1	51.2 $\pm$ 11.8	0.22 ($P=0.82$)	—
Post-test (%)	62.4 $\pm$ 11.3	72.1 $\pm$ 10.7	5.51 ($P<0.001$)	0.54
Post-test gain (%)	11.6 $\pm$ 9.4	20.9 $\pm$ 10.1	6.01 ($P<0.001$)	0.72
Retention score (%)	55.9 $\pm$ 12.4	67.3 $\pm$ 11.2	6.06 ($P<0.001$)	0.66
Retention–post-test gap	–6.5 $\pm$ 8.3	–4.8 $\pm$ 7.6	1.37 ($P=0.17$)	—

The blended-FA cohort outperformed controls on the DLAT post-test by 9.7 percentage points ($t(174)=5.51$, $P<0.001$, $d=0.54$), a medium effect consistent with meta-analytic estimates for formative assessment interventions. The post-test gain for the blended-FA cohort ($M=20.9\%$) was nearly 80% larger than that of controls ($M=11.6\%$), with a large between-group effect ($d=0.72$). At the eight-week retention test, the blended-FA cohort maintained a mean score of 67.3% compared with 55.9% in the control cohort ($d=0.66$). The retention–post-test gap did not differ significantly between conditions ($P=0.17$), indicating that the blended-FA cohort entered the retention interval with a higher absolute score, but the rate of forgetting over eight weeks was comparable between groups.

Because the primary aim was to isolate the intervention effect from any residual baseline variation, ANCOVA provided the confirmatory analysis (**Table 4**). After adjusting for pre-test performance, the effect of condition on the DLAT post-test remained highly significant ($F(1, 173)=29.84$, $P<0.001$, partial $\eta^2=0.147$), a large adjusted effect, and it persisted at the eight-week retention test ($F(1, 173)=36.72$, $P<0.001$, partial $\eta^2=0.175$). Adjusted post-test means were 71.9% for the blended-FA cohort versus 62.6%

for controls, closely matching the unadjusted values and confirming that the advantage was not an artifact of baseline variation.

Table 4. ANCOVA of post-test outcomes with the corresponding pre-test score as covariate

Outcome	Adj. mean (Control)	Adj. mean (Blended-FA)	<i>F</i> (1,173)	Partial η^2
DLAT post-test (%)	62.6	71.9	29.84 ($P<0.001$)	0.147
DLAT retention (%)	56.1	67.1	36.72 ($P<0.001$)	0.175
DSSE (1–5)	3.49	4.11	33.91 ($P<0.001$)	0.164

4.2. Course grade outcomes

Final grade distributions differed significantly between conditions ($\chi^2(4)=14.23$, $P=0.007$). **Table 5** shows the distribution across five grade bands; **Figure 2(c)** compares failure and A-grade rates.

Table 5. Final grade distribution by condition (%)

Grade band	Control (%)	Blended-FA (%)
A (≥ 85)	15.4	28.6
B (75–84)	28.2	37.1
C (65–74)	34.6	25.7
D (55–64)	8.9	3.8
F (< 55)	13.2	4.8

The failure rate in the blended-FA cohort (4.8%) was less than half that in the control cohort (13.2%), while the proportion of A-grade students was nearly double (28.6% vs. 15.4%). The mean final course score was 72.8 ± 9.1 for the blended-FA cohort and 68.3 ± 10.2 for controls ($t(174)=3.02$, $P=0.003$, $d=0.44$). These distributional shifts suggest that the intervention both raised the ceiling for high achievers—through Bloom Level 5–6 quiz items and peer critique tasks—and raised the floor for struggling students by providing timely corrective feedback that prevented the accumulation of uncorrected misconceptions.

4.3. Self-efficacy outcomes (RQ2)

Pre-intervention DSSE scores did not differ between conditions ($M=3.21$ vs. 3.19 ; $t(174)=0.18$, $P=0.86$). At post-intervention, the blended-FA cohort reported a substantially higher mean DSSE score ($M=4.12$ vs. 3.48 ; $t(174)=5.82$, $P<0.001$, $d=0.67$). An ANCOVA adjusting for pre-intervention DSSE confirmed the condition effect ($F(1,173)=33.91$, $P<0.001$, partial $\eta^2=0.164$; **Table 4**). **Figure 2(b)** displays these results. Gains in DSSE correlated positively with DLAT post-test gains within the blended-FA cohort ($r=0.41$, $P<0.001$), suggesting that self-efficacy growth and knowledge consolidation were mutually reinforcing rather than independent outcomes.

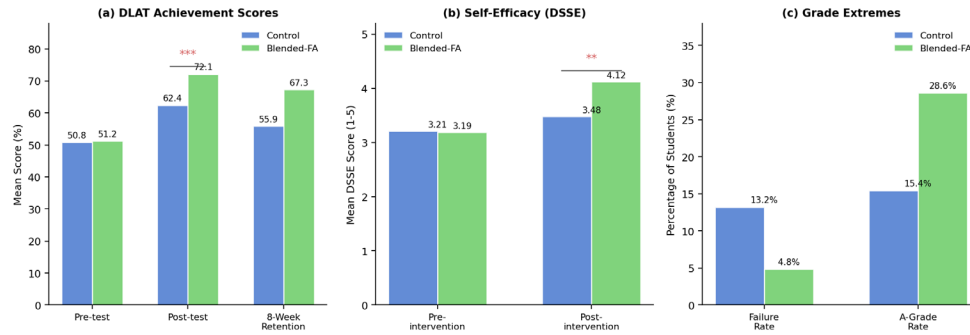


Figure 2. Comparison of outcomes between the control and blended-FA cohorts. (a) DLAT achievement scores at pre-test, post-test, and eight-week retention. (b) Domain-Specific Self-Efficacy (DSSE) scores before and after the intervention. (c) Failure rate and A-grade rate by condition. $**P < 0.01$; $***P < 0.001$.

4.4. Interview findings: mechanisms of learning

Thematic analysis yielded three primary themes—the quiz as a reality check, productive confusion from prediction tasks, and teaching-to-learn through peer critique —each mapping coherently to one intervention component. Representative quotations are presented in **Table 6**.

Table 6. Interview themes and representative student quotations

Theme	Representative quotation	Linked component
Quiz as a reality check	“Every Sunday I thought I understood the week’s content, but the quiz showed me exactly which part I hadn’t actually grasped. The feedback told me why I was wrong, so I fixed it before moving on.” — Blended-FA student S07	Weekly online quizzes
Productive confusion	“When the lab result didn’t match my prediction, I had to work out why. That process of resolving the mismatch was when I really understood propagation delay.” — Blended-FA student S14	Pre-lab prediction tasks
Teaching to learn	“Writing the critique forced me to articulate things I thought I understood but couldn’t explain. Explaining to my peer made it stick in a way that solving problems alone never did.” — Blended-FA student S03	Peer critique sessions

Control-group interviewees most frequently described a pattern of intensive study concentrated in the two weeks before each examination, with little structured engagement between those peaks. Several reported discovering the depth of their conceptual gaps only when attempting past-paper questions immediately before the final examination—when corrective time was limited. This bunching pattern mirrors the “illusion of knowing” documented in the metacognition literature ^[24] and represents precisely the failure mode that distributed formative assessment is designed to pre-empt.

5. Discussion

This study provides quasi-experimental evidence that embedding three structurally distinct formative assessment activities into a Digital Logic Design course was associated with meaningful improvements in short-term learning achievement ($d=0.54$), eight-week retention ($d=0.66$), and domain-specific self-efficacy ($d=0.67$), alongside a reduction of approximately two-thirds in the course failure rate. These effect sizes are consistent with—and at the upper end of—the range reported in domain-general formative assessment meta-analyses ^[7,8], suggesting that digital logic design is a responsive domain for such interventions, perhaps because the prevalence of precise, verifiable answers makes automated feedback particularly feasible and

credible.

Positioning this contribution against prior digital logic teaching reforms clarifies its distinctive focus (**Table 7**). Whereas earlier reform efforts have concentrated on restructuring the laboratory—most prominently through FPGA-based project work—and have evaluated success primarily through laboratory or skill-based scores, the present study intervenes at the level of routine formative assessment and evaluates a dual outcome set encompassing both conceptual achievement and learner self-efficacy. The approach also requires no additional hardware or curriculum redesign, lowering the barrier to adoption.

Table 7. Comparison of the present study with representative prior digital logic teaching reforms

Study/approach	Primary method	Outcome measured	Gap addressed here
FPGA-based laboratory reform ^[20]	Hardware/project restructuring	Laboratory and skill scores	Requires new hardware; conceptual gains not isolated
Competency-based curriculum design ^[19]	Curriculum mapping to competencies	Program-level competency attainment	Coarse-grained; no within-semester feedback loop
Active learning/peer instruction ^[16,17]	In-class response and discussion	Examination performance	Motivational outcomes rarely measured
This study (blended embedded FA)	Weekly quizzes + prediction + peer critique	Achievement, retention, and self-efficacy	No new hardware; dual cognitive-motivational gains

The self-efficacy findings deserve particular emphasis. An increase of $d=0.67$ on the DSSE, together with its positive correlation with DLAT gain ($r=0.41$), indicates that the blended-FA model was not merely producing knowledge acquisition while leaving motivational structure unchanged, but was actively reshaping students' beliefs about their own competence. This is important because self-efficacy predicts not only current-semester achievement but also engagement in subsequent courses and willingness to pursue careers in technically demanding fields^[11,12]. A course intervention that builds both knowledge and confidence may therefore have downstream effects well beyond the measured outcomes.

The qualitative data indicate that each intervention component engaged a distinct cognitive mechanism. The weekly quiz cycle leveraged the testing effect: retrieval practice enhances long-term retention compared with equivalent restudying time^[5,6]. Pre-lab prediction tasks engaged anticipatory prediction error^[25]: when an explicit prediction is disconfirmed by simulation or hardware behavior, the resulting surprise triggers deeper processing of the discrepant information than would occur if the student simply observed the outcome without prior commitment. Peer critique sessions created opportunities for elaborative explanation—the generation of causal accounts of circuit behavior—which consistently promotes conceptual understanding beyond what procedural practice alone achieves^[18,26]. The convergence of three complementary mechanisms, each targeting a different aspect of durable learning, likely explains why effects were not limited to a single outcome domain.

5.1. Limitations

Several limitations qualify these findings. First, and most fundamentally, the quasi-experimental design assigned intact cohorts to conditions across consecutive semesters rather than randomizing individual students, raising the possibility of selection bias. Although the groups were statistically equivalent on age, prior GPA, and DLAT pre-test scores, and although ANCOVA adjusted for baseline performance, unmeasured differences—such as cohort-level motivation, concurrent coursework load, or admission-year variation—cannot be fully excluded and may have contributed to the observed effects.

Second, all participants were drawn from a single research-intensive institution with a competitive admission profile. This single-institution design limits external validity: the magnitude of the effects—and possibly their direction—might differ at institutions serving students with weaker mathematical preparation, in substantially larger classes, or under different resourcing for feedback provision. Multi-site replication is required before the model can be recommended for broad adoption.

Third, the self-efficacy outcome relied on a self-report instrument (the DSSE). Self-report measures are susceptible to social-desirability and demand characteristics, and students in a visibly novel instructional condition may have reported inflated confidence to align with perceived instructor expectations. Although the convergence of self-efficacy gains with objective achievement gains partially mitigates this concern, future work should incorporate behavioral indices to triangulate the self-report data.

Fourth, the study is vulnerable to a Hawthorne effect: the blended-FA cohort received demonstrably more instructional attention—more frequent feedback, more structured activities, and the implicit signal of participating in an innovation—than the control cohort. Part of the observed benefit may therefore stem from heightened engagement and attention *per se* rather than from the specific formative assessment mechanisms. A stronger future design would equate time-on-task and instructor contact across conditions, for example by providing the control group with an equivalent volume of alternative, non-formative activity.

Finally, the three components were delivered as an integrated bundle, precluding attribution of the effects to any single element. A future factorial design that systematically varies the presence of each component would identify the minimal effective combination for instructors operating under constrained time and capacity.

A practical note concerns the interval between data collection (the 2023–2024 academic year) and publication. This gap reflects the time required for the eight-week retention follow-up, verbatim transcription and reflexive coding of the interview data, and successive rounds of analysis and peer review; it does not affect the internal validity of the reported within-study comparisons.

6. Conclusion

This study demonstrated that a blended instructional model embedding three types of formative assessment activity into a Digital Logic Design course produced significant and practically meaningful improvements in learning achievement, long-term retention, and learner self-efficacy relative to conventional lecture-and-examination instruction. The intervention reduced the course failure rate from 13.2% to 4.8% and increased the A-grade proportion from 15.4% to 28.6%, while building students' confidence in their circuit-design competence. The self-efficacy gains may have positive downstream effects on engagement in subsequent advanced courses.

For practitioners, the study offers a replicable model requiring no curriculum redesign, new laboratory hardware, or substantial increases in instructor workload. Weekly online quizzes with automated feedback can be built in any standard learning management system; pre-lab prediction tasks require no additional class time; and peer critique sessions occupy three structured one-hour slots per semester. Future research should examine whether the individual components contribute independently to the observed effects, how the model performs with weaker student populations and under larger class sizes, and whether the self-efficacy gains observed here translate into improved performance and persistence in subsequent advanced courses.

Funding

This work was supported by the State Key Laboratory of Processors, Institute of Computing Technology, Chinese Academy of Sciences, under Open Project Grant No. CLQ202516, and by the Fundamental Research Funds for the Central Universities (Research on Key Technologies for Post-Quantum Cryptographic Chip Design and Quantum Circuit Evaluation, Grant No. 3282026039).

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Mansouri M, 2026, Active Learning in Engineering Education: A Case Study from a Digital Logic Design course. *Applied Sciences*, 16(6): 2880. <https://doi.org/10.3390/app16062880>
- [2] Theobald EJ, Hill MJ, Tran E, et al., 2020, Active Learning Narrows Achievement Gaps for Underrepresented Students in Undergraduate Science, Technology, engineering, and Math. *Proceedings of the National Academy of Sciences*, 117(12): 6476–6483. <https://doi.org/10.1073/pnas.1916903117>
- [3] Sweller J, 2023, The Development of Cognitive Load Theory: Replication Crises and Incorporation of Other Theories Can Lead to Theory Expansion. *Educational Psychology Review*, 35(4): 95. <https://doi.org/10.1007/s10648-023-09817-2>
- [4] Morris R, Perry T, Wardle L, 2021, Formative Assessment and Feedback for Learning in Higher Education: A Systematic Review. *Review of Education*, 9(3): e3292. <https://doi.org/10.1002/rev3.3292>
- [5] Yang C, Luo L, Vadillo MA, et al., 2021, Testing (Quizzing) Boosts Classroom Learning: A Systematic and Meta-Analytic Review. *Psychological Bulletin*, 147(4): 399–435. <https://doi.org/10.1037/bul0000309>
- [6] Agarwal PK, Nunes LD, Blunt JR, 2021, Retrieval Practice Consistently Benefits Student Learning: A Systematic Review of Applied Research in Schools and Classrooms. *Educational Psychology Review*, 33(4): 1409–1453. <https://doi.org/10.1007/s10648-021-09595-9>
- [7] Wisniewski B, Zierer K, Hattie J, 2020, The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10: 3087. <https://doi.org/10.3389/fpsyg.2019.03087>
- [8] McDermott KB, 2021, Practicing Retrieval Facilitates Learning. *Annual Review of Psychology*, 72: 609–633. <https://doi.org/10.1146/annurev-psych-010419-051019>
- [9] Bandura A, 1997, *Self-Efficacy: The Exercise of Control*. W. H. Freeman, New York.
- [10] Schunk DH, DiBenedetto MK, 2020, Motivation and Social Cognitive Theory. *Contemporary Educational Psychology*, 60: 101832. <https://doi.org/10.1016/j.cedpsych.2019.101832>
- [11] Chen X-Y, Usher EL, Roeder M, et al., 2023, Mastery, Models, Messengers, and Mixed Emotions: Examining the Development of Engineering Self-Efficacy by gender. *Journal of Engineering Education*, 112(1): 64–89. <https://doi.org/10.1002/jee.20494>
- [12] Lee Y-K, Freer E, Robinson KA, et al., 2022, The Multiplicative Function of Expectancy and Value in Predicting Engineering Students' Choice, Persistence, and Performance. *Journal of Engineering Education*, 111(3): 531–553. <https://doi.org/10.1002/jee.20456>
- [13] Cao W, 2023, A Meta-Analysis of Effects of Blended Learning on Performance, Attitude, Achievement, and Engagement Across Different Countries. *Frontiers in Psychology*, 14: 1212056. <https://doi.org/10.3389/>

- [14] Double KS, McGrane JA, Hopfenbeck TN, 2020, The Impact of Peer Assessment on Academic Performance: A Meta-Analysis of Control Group Studies. *Educational Psychology Review*, 32(2): 481–509. <https://doi.org/10.1007/s10648-019-09510-3>
- [15] Ting FST, Shroff RH, Lam WH, et al., 2022, A Meta-Analysis of Studies on the Effects of Active Learning on Asian Students' Performance in Science, Technology, Engineering and Mathematics (STEM) Subjects. *The Asia-Pacific Education Researcher*, 32(3): 379–400. <https://doi.org/10.1007/s40299-022-00661-6>
- [16] Nguyen KA, Borrego M, Finelli CJ, et al., 2021, Instructor Strategies to Aid Implementation of Active Learning: A Systematic Literature Review. *International Journal of STEM Education*, 8: 9. <https://doi.org/10.1186/s40594-021-00270-7>
- [17] Fan H, Xie H, Feng Q, et al., 2022, Interdisciplinary Project-Based Learning: Experiences and Reflections from Teaching Electronic Engineering in China. *IEEE Transactions on Education*, 66(1): 73–82. <https://doi.org/10.1109/TE.2022.3186184>
- [18] Corso Pinzón AF, Ramírez-Echeverry JJ, Restrepo-Calle F, 2024, Automated Grading Software Tool with Feedback Process to Support Learning of Hardware Description Languages. *Research and Practice in Technology Enhanced Learning*, 19: 015. <https://doi.org/10.58459/rptel.2024.19015>
- [19] Tian Y, Yu J, 2026, Innovation and Teaching Reform of the Talent Training Mode in Integrated Circuit Science and Engineering under the Background of New Engineering: A Case Study of North University of China. *Journal of Contemporary Educational Research*, 10(7): 1–9.
- [20] Sabanwar V, Trimukhe P, Gudla A, et al., 2023, Improving Learning Experience to Teach FPGA Design at Scale. In: *Proceedings of the 16th Annual ACM India Compute Conference (COMPUTE 2023)*. ACM, New York, 31–36. <https://doi.org/10.1145/3627217.3627225>
- [21] Miller MD, Linn RL, Gronlund NE, 2013, *Measurement and Assessment in Teaching* (11th edn). Pearson, Boston.
- [22] Pajares F, Miller MD, 1994, Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis. *Journal of Educational Psychology*, 86(2): 193–203. <https://doi.org/10.1037/0022-0663.86.2.193>
- [23] Braun V, Clarke V, 2022, Conceptual and Design Thinking for Thematic Analysis. *Qualitative Psychology*, 9(1): 3–26. <https://doi.org/10.1037/qup0000196>
- [24] Koriati A, Bjork RA, 2005, Illusions of Competence in Monitoring One's Knowledge During Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2): 187–194. <https://doi.org/10.1037/0278-7393.31.2.187>
- [25] Brod G, Hasselhorn M, Bunge SA, 2018, When Generating a Prediction Boosts Learning: The Element of Surprise. *Learning and Instruction*, 55: 22–31. <https://doi.org/10.1016/j.learninstruc.2018.01.013>
- [26] Chi MTH, Wylie R, 2014, The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4): 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- [27] Yu Q, Yu K, Li B, Wang Q, 2023, Effectiveness of Blended Learning on Students' Learning Performance: A Meta-Analysis. *Journal of Research on Technology in Education*, 57(3): 499–520. <https://doi.org/10.1080/15391523.2023.2264984>

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.