International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 7, pp. 514-534, July 2026
https://doi.org/10.26803/ijlter.25.7.23
Received May 16, 2026; Revised Jun 19, 2026; Accepted Jun 22, 2026

Peer Feedback Training in L2 English Writing: A Systematic Literature Review of Empirical Studies from 2016 to 2025

Jing Wang
Universiti Kebangsaan Malaysia, Selangor, Malaysia
Zhengzhou Business University, Zhengzhou, Henan, China

Melor Md Yunus* **and Nur Ainil Sulaiman**
Universiti Kebangsaan Malaysia, Selangor, Malaysia

Abstract. Although peer feedback has been widely used in second language (L2) English writing instruction, its pedagogical value is often constrained by vague comments, limited understanding of evaluation criteria, low learner confidence, weak peer trust, and insufficient uptake of revision. Guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses framework, this systematic literature review identified and synthesized empirical studies published from 2016 to 2025 on peer feedback training in English as a Foreign Language and English as a Second Language writing contexts. Database searches in Scopus, Educational Resources Information Center, and Web of Science were followed by deduplication, title and abstract screening, full-text eligibility assessment, and quality appraisal. Nineteen empirical studies were included in a narrative thematic synthesis focusing on instructional design, reported learning outcomes, and implementation challenges. The synthesis suggests that peer feedback is more effective when supported by rubrics or checklists, teacher modeling, guided peer-review tools, peer dialogue, AI-supported scaffolding, and revision-reflection tasks. Reported outcomes included improvements in feedback quality, revision uptake, writing performance, feedback literacy, motivation, self-efficacy, engagement, and self-regulated writing. Recurring implementation challenges included uneven learner readiness, limited feedback specificity, insufficient peer trust, time constraints, and the need for teacher monitoring and systematic guidance. Based on these findings, the review proposes a lightweight peer feedback training framework for L2 writing instruction and future intervention research.

Citation:
Wang, J., Yunus, M., Sulaiman, N.A. (2026). Peer Feedback Training in L2 English Writing: A Systematic Literature Review of Empirical Studies from 2016 to 2025. *International Journal of Learning, Teaching and Educational Research, 25(7),* 514–534. https://doi.org/10.26803/ijlter.25.7.23

Corresponding author: Melor Md Yunus; melor@ukm.edu.my*

© The Authors
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Keywords: EFL writing; feedback literacy; L2 English writing; peer feedback training; systematic literature review

1. Introduction

Second language (L2) English writing, including English as a Foreign Language (EFL) and English as a Second Language (ESL) writing, requires learners to coordinate linguistic accuracy, rhetorical organization, argument development, coherence, audience awareness, and revision decisions. Feedback is therefore central to writing instruction because it helps learners identify gaps between current and expected performance and supports revision-oriented learning (Hattie & Timperley, 2007; Hyland & Hyland, 2006). In higher education EFL and ESL contexts, teacher feedback remains essential, but it is difficult to provide repeated and detailed responses to every student, particularly in large or writing-intensive classes. This practical constraint has encouraged researchers and teachers to examine peer feedback as a scalable and learner-centered complement to teacher feedback.

Peer feedback can increase the amount of feedback students receive, expose writers to multiple readers, and develop students' evaluative judgment as reviewers. However, peer feedback is not inherently effective: students may offer vague praise, focus mainly on grammar or vocabulary, avoid critical comments for affective reasons, distrust peer input, or fail to use comments during revision. These concerns are particularly relevant in L2 English writing classrooms, where learners may lack confidence in both their own writing and their ability to evaluate peers' texts. For this reason, the present review treats peer feedback as a trained and scaffolded literacy practice rather than as an unguided classroom activity.

In practice, higher education writing teachers need manageable ways to support peer review without adding excessive workload, while in research, previous studies have examined peer feedback in second language writing, but less attention has been paid to the training components that make it useful, trustworthy, and revision oriented. These components may include rubric explanation, sample-text calibration, teacher modeling, guided review sheets, feedback language support, peer dialogue, online tools, AI-supported scaffolding, and revision-reflection tasks.

Three gaps motivate this review. First, existing studies use varied training designs, making it difficult to identify common instructional features. Second, reported outcomes are fragmented across feedback quality, feedback literacy capacities, revision uptake, writing performance, motivation, self-efficacy, and learner perceptions. Third, implementation conditions such as teacher monitoring, learner readiness, technology access, peer trust, group participation, and time allocation remain inconsistently examined. This review therefore synthesizes empirical studies from 2016 to 2025 concerning instructional design, learning outcomes, and pedagogical implementation.

This systematic literature review synthesizes empirical studies on peer feedback training in second language English writing in higher education, with a focus on instructional design, reported learning outcomes, and pedagogical challenges affecting classroom implementation. It addresses the following research questions:
RQ1: What instructional designs of peer feedback training have been used in L2 English writing instruction?
RQ2: What learning outcomes have been reported regarding feedback quality, feedback literacy capacities, revision uptake, and writing performance?
RQ3: What pedagogical challenges and implementation conditions influence the effectiveness of peer feedback training in writing classrooms?

2. Literature Review

2.1 Peer Feedback and L2 English Writing

Peer feedback refers to a writing-related instructional process in which learners review peers' drafts, provide comments, evaluate suggestions, and use feedback to support revision. In L2 English writing classrooms, peer feedback has pedagogical value because it can increase feedback opportunities, expose writers to multiple readers, and help students develop evaluative judgment as both writers and reviewers. Compared with teacher-only feedback, peer feedback may also encourage students to become more actively involved in identifying textual problems, discussing revision options, and reflecting on writing quality.

However, peer feedback in L2 writing is not automatically effective as students may provide vague praise, focus narrowly on grammar and vocabulary, avoid critical comments, distrust peer suggestions, or fail to incorporate feedback into revision. These problems are especially salient in EFL and ESL contexts, where learners may have limited confidence in their linguistic knowledge and limited experience using criteria to evaluate writing. Therefore, the effectiveness of peer feedback depends not only on the exchange of comments but also on whether students understand evaluation criteria, generate specific and constructive feedback, engage in peer dialogue, and make informed revision decisions.

For these reasons, recent research increasingly treats peer feedback as a trained and scaffolded literacy practice. Systematic guidance, such as rubric familiarization, teacher modeling, guided review sheets, feedback language support, peer discussion, technology-supported review, and revision-reflection tasks, can help students move beyond surface-level correction toward more meaningful feedback use. From this perspective, peer feedback training becomes the primary site for developing the feedback literacy capacities explored in the following section.

2.2 Peer Feedback Training and Feedback Literacy

Peer feedback training refers to instructional preparation that helps learners understand writing criteria, generate useful comments, interact with peers, evaluate feedback, and revise texts. Such training may include rubric familiarization, teacher modeling, sample-comment analysis, guided peer-review sheets, and feedback language frames. It may also involve peer dialogue,

technology-supported review, AI-supported scaffolding, and revision-reflection tasks. These components are closely related to student feedback literacy, which involves appreciating feedback, making judgments, managing affect, and taking action on feedback (Carless & Boud, 2018; Winstone et al., 2017).

In L2 writing classrooms, feedback literacy is particularly important because students must interpret both language-level and meaning-level comments and decide which suggestions can improve the rhetorical, organizational, and linguistic quality of their drafts. Peer feedback training can therefore be viewed as an instructional route through which learners move from simple error correction toward more dialogic, evaluative, and revision-oriented feedback practice.

2.3 Existing Reviews and Remaining Research Gap

Existing reviews have provided important evidence on peer feedback in second language writing and technology-supported English language education. Yu and Lee (2016) synthesized peer feedback research in second language writing more broadly, while recent reviews have examined ICT-supported academic English writing (Yingbao et al., 2025), techno-pedagogical integration in ESL classrooms (Leong et al., 2024), blended ESL learning (Ramalingam et al., 2022), and social networking sites in English language education (Tan & Yunus, 2023). However, prior reviews have tended either to address peer feedback at a broad level or to focus on technology-enhanced language learning more generally. They have not specifically synthesized how peer feedback training is designed, what learning outcomes are reported, and what implementation conditions shape its effectiveness in higher education L2 English writing contexts.

To address this gap, the present review synthesizes empirical studies published from 2016 to 2025 through the lens of three dimensions: instructional design, learning outcomes, and pedagogical implementation. By limiting the scope to higher education L2/EFL/ESL English writing, the review provides a focused account of how peer feedback training can be designed and implemented in classroom-based writing instruction.

3. Methods

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, 2020) guidelines throughout the literature search, screening, eligibility assessment, and final study inclusion process (Page et al., 2021). Because the included studies differed substantially in instructional interventions, outcome measures, participant groups, and analytical designs, this review adopted narrative thematic synthesis rather than quantitative meta-analysis. The data set for this analysis included only original empirical studies and excluded theoretical discussions, existing reviews, textbook chapters, and practice-oriented articles. The complete research process of this systematic review is shown in Figure 1. The review protocol was designed to ensure transparency, reproducibility, and conceptual comparability across the included studies.

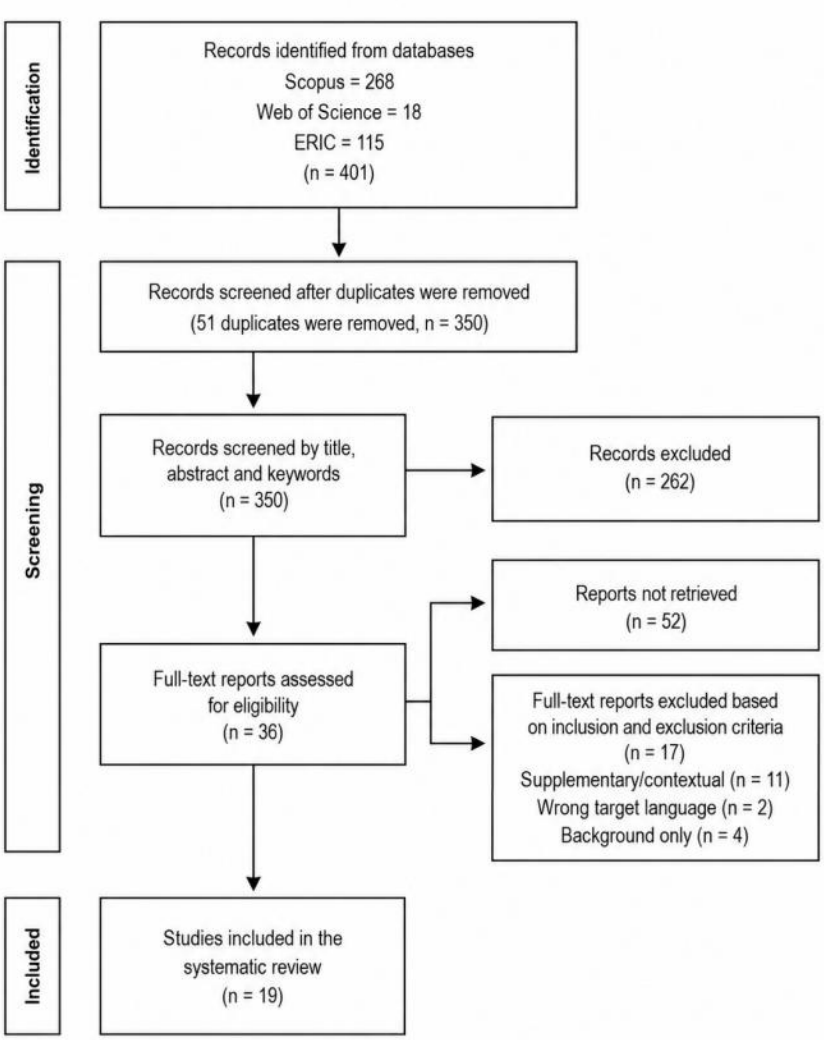

Figure 1: PRISMA flow diagram of the study selection process

3.1 Identification Phase

Relevant records were retrieved from three major databases: Scopus, Web of Science, and ERIC (Educational Resources Information Center). These databases were selected to balance broad coverage and educational relevance: Scopus covers applied linguistics, general education, educational technology, and higher education; Web of Science includes peer-reviewed journals in language education and educational research; and ERIC focuses on education, classroom practice, and teaching methods. Google Scholar was not used as a primary search database because its search results are less stable, its replication procedures are less transparent, and its batch export functions may complicate consistent record management.

Before duplicate removal, 401 records were retrieved: 268 from Scopus, 115 from ERIC, and 18 from Web of Science. After 51 duplicate records were removed, 350 records proceeded to title and abstract screening. The search string combined three conceptual categories: peer feedback-related terms, training or scaffolding-

related terms, and terms related to L2/EFL/ESL writing. Terms such as scoring criteria and feedback literacy were not used as mandatory search terms but were instead applied later as coding categories to avoid making the initial search scope too narrow.

Table 1: Search log and record sources

Database	Records	Export/search date	Note
Scopus	268	2026-03-06	Database export after applying the 2016-2025, English language and article-related filters.
ERIC	115	2026-04-07	ERIC export; records outside the 2016-2025 range were removed during cleaning.
Web of Science	18	2026-04-08	Re-search for and import of English language article records within the final date range.
Total	401	April 2026	Total records before deduplication.

The database records in Table 1 were retrieved using the search strings summarized in Table 2.

Table 2: Core search strings used in the database searches

Database	Search string
Scopus	TITLE-ABS-KEY (("peer feedback" OR "peer review" OR "peer response" OR "peer assessment") AND (training OR guided OR scaffold* OR instruction OR intervention OR structured) AND (writing OR "L2 writing" OR "EFL writing" OR "ESL writing") AND (EFL OR ESL OR L2 OR "second language" OR "foreign language"))
Web of Science	TS=(("peer feedback" OR "peer review" OR "peer response" OR "peer assessment") AND (training OR guided OR scaffold* OR instruction OR intervention OR structured) AND (writing OR "L2 writing" OR "EFL writing" OR "ESL writing") AND (EFL OR ESL OR L2 OR "second language" OR "foreign language"))
ERIC	("peer feedback" OR "peer review" OR "peer response" OR "peer assessment") AND (training OR guided OR scaffolded OR instruction OR intervention OR structured) AND (writing OR "EFL writing" OR "ESL writing" OR "L2 writing")

This targeted search was designed to retrieve literature on peer feedback and structured scaffolding interventions in L2 English writing while avoiding an overly narrow initial search scope.

3.2 Screening Phase

The inclusion criteria reflect the research scope of this review: the educational setting is higher education; the study type is original empirical research; the

writing context is second language, EFL, or ESL English; and the instructional content involves peer feedback training or structured scaffolding. These restrictions were used to keep the evidence base comparable across educational level, target language, study type, and instructional focus. Supplementary or scope-dependent studies were excluded from the main synthesis when they did not fully meet the higher education, empirical-study, or direct L2 English writing peer feedback training criteria.

Table 3 presents the inclusion and exclusion criteria, which were developed in line with the review aim, PRISMA guidance, and the need to ensure comparability across publication type, date range, language, context, peer feedback focus, training/scaffolding features, and reported findings.

Table 3: Inclusion and exclusion criteria

Criterion	Inclusion criteria	Exclusion criteria
Publication type	Peer-reviewed empirical studies.	Reviews, conceptual papers, editorials, theses, books, handbook chapters, reports, or practice-only articles.
Date and language	2016-2025; English language publications.	Outside the year range or non-English publications.
Context	Higher education L2/EFL/ESL English writing.	School-age contexts, non-English target language writing, or non-writing contexts.
Peer feedback focus	Peer feedback, peer review, peer response, or peer assessment in writing.	Teacher-only, automated-only, or self-assessment-only studies.
Training/scaffolding	Training, guidance, scaffolding, rubrics, criteria, checklists, structured prompts, or technology- or AI-supported scaffolding.	Unguided peer feedback with no training, guidance, or structure.
Findings	Reports outcomes, perceptions, revision evidence, feedback quality, or implementation findings.	Insufficient method or results information.

These criteria were applied during both title/abstract screening and full-text eligibility assessment. Records were retained only when they aligned with the target context of higher education L2/EFL/ESL English writing and included peer feedback training, guidance, scaffolding, rubrics, criteria, structured procedures, or technology-/AI-supported support as part of the peer feedback process.

All retrieved records were exported to a screening workbook. Duplicate records were removed through DOI/title matching and manual checking. Title and abstract screening classified records as Include, Unclear, or Exclude, and the preliminary Include and Unclear records were treated as reports for retrieval. Of

the 88 reports for retrieval, 52 could not be accessed as full-text reports after final retrieval attempts, leaving 36 full-text reports for eligibility assessment. Under the strict higher education and empirical-study criteria, 19 studies met all criteria for the main thematic synthesis. The remaining full-text reports were classified as supplementary or scope-dependent evidence, wrong target language studies, or background-only sources. The final screening numbers are reported in Table 4.

Table 4: Final PRISMA counts for the study selection process

Stage	Item	Count	Notes
Identification	Records identified from databases	401	Scopus 268; ERIC 115; Web of Science 18.
Screening	Duplicates removed	51	DOI/title matching and manual checking.
Screening	Records screened	350	Title/abstract screening.
Screening	Records excluded at title/abstract stage	262	Manual preliminary exclusion.
Retrieval	Reports sought for retrieval	88	Preliminary Include + Unclear.
Retrieval	Reports not retrieved	52	Final retrieval outcome: these reports could not be accessed as full-text reports after final retrieval attempts.
Eligibility	Full-text reports assessed	36	Final full-text eligibility set.
Eligibility	Reports excluded from the core synthesis	17	11 supplementary/scope-dependent; 2 wrong target language; 4 background only.
Included	Studies included in the main synthesis	19	Higher education primary empirical studies.

Data retrieval transparency and potential selection bias were considered during screening. The 52 unretrieved reports were records for which full texts could not be obtained after the final retrieval attempts, usually because the database entry provided only bibliographic information or abstract-level access, institutional full-text access was unavailable, or the report was not publicly accessible at the time of screening. This relatively high number may introduce a degree of selection bias because unavailable reports could contain different methods or findings. To reduce this risk, the review reports the retrieval process transparently and limits the core synthesis to studies for which full-text evidence could be examined.

Quality appraisal examined whether each study had a clear objective, an appropriate L2 English writing context, a peer feedback component, training or structured scaffolding, clear methods, and reportable outcomes. The initial appraisal was conducted by the first author, and ambiguous cases were rechecked against the eligibility criteria and full-text evidence before final classification. This procedure was used not to rank studies by effect size but to determine whether each study provided sufficient methodological and contextual information for the

thematic synthesis. Studies scoring 4–5 were considered suitable for the core synthesis; studies with moderate relevance or scope limitations were retained only as supplementary/contextual evidence; and studies below the threshold or outside the target language/context were excluded from the core synthesis.

Quality appraisal used five criteria: clear research objectives; direct relevance to peer feedback/review/response/assessment in writing; presence of training/guidance/scaffolding/rubrics/criteria/structured procedures; clear reporting of outcomes or implementation findings; and sufficient methodological description. Each criterion was scored as 0, 0.5, or 1. Studies scoring 4–5 were included in the core synthesis; studies scoring 3–3.5 or having scope limitations were retained only as supplementary/contextual evidence; and studies below 3 or outside the target language/context were excluded from the core synthesis.

3.3 Inclusion Phase

Following identification, screening, full-text eligibility assessment, and quality appraisal, 19 higher education empirical studies were included in the main thematic synthesis. The 11 supplementary studies were excluded from the main synthesis because the confirmed scope was limited to higher education empirical studies directly related to L2 English writing peer feedback training or structured support. Table 5 summarizes the 19 studies that formed the main evidence base for the thematic synthesis.

Table 5: Summary of the 19 core studies included in the main thematic synthesis

Study	Context/ participants	Design and support	Main outcome or finding
Deng & Sitthitikul (2025)	Chinese EFL university students	Mixed-methods quasi-experimental; guided dialogic peer feedback in blended writing	Improved argumentative writing performance and positive perceptions
Do (2024)	Vietnamese EFL college students	Qualitative case study; Google Docs/Zoom-based e-peer feedback training	Supported more purposeful online peer feedback among inexperienced learners
Lin & Hwang (2025)	Taiwanese university EFL students	Quasi-experimental; AI-supported procedural scaffolding in flipped peer assessment	Improved writing accuracy and higher-order thinking indicators
Sun & Wang (2025)	Chinese non-English-major graduate students	Quasi-experimental; teacher intervention with peer feedback	Improved revision uptake, accurate revision, and self-correction

Study	Context/ participants	Design and support	Main outcome or finding
Liu & Luo (2025)	Chinese undergraduate English majors	Mixed-methods online peer feedback intervention with criteria, group review, discussion, and revisions	Enhanced perceived peer feedback literacy
Zheng (2025)	Chinese EFL business freshmen	PeerMark-based technology-enhanced peer feedback workshops	Supported revision motivation and revision quality
Ghaneiarani et al. (2024)	Iranian intermediate EFL learners	Experimental/quasi-experimental; peer vs. teacher feedback in learning-oriented assessment	Peer feedback supported writing gains and retention
Hanh & Tinh (2022)	Vietnamese non-English-major university students	Quasi-experimental mixed-methods; peer-review checklist	Improved task fulfillment, vocabulary, and student perceptions
Awada & Diab (2023)	Lebanese EFL university learners	Experimental comparison; online written vs. face-to-face oral peer review	Online peer review produced stronger argumentative synthesis writing outcomes
Min (2016)	Taiwanese college freshmen	Peer review training experiment; teacher modeling and feedback	Improved trained peer-review skills
Cui et al. (2021)	Chinese EFL university students	Between-groups comparison; trained peer vs. teacher feedback	Produced writing and self-efficacy gains comparable to teacher feedback
Guo et al. (2024)	Chinese undergraduate EFL students	Quasi-experimental; AI-supported peer feedback through EvaluMate	Improved feedback quality and feedback providers' writing ability
Weng et al. (2024)	Macau Chinese EFL English majors	Mixed-methods intervention; peer vs. teacher feedback activities	Improved writing feedback literacy
Hassani et al. (2023)	Iranian International English Language Testing System candidates	Experimental comparison; synchronous/asynchronous computer-mediated peer feedback	Computer-mediated peer feedback supported writing development
Zhang & Yu (2022)	Chinese university L2 writers	Qualitative case study; systemic functional linguistics (SFL)-based meaning-making peer feedback training	Shifted feedback from structural accuracy toward meaning making

Study	Context/ participants	Design and support	Main outcome or finding
Guo et al. (2025)	Chinese undergraduates in Hong Kong/Macau context	Mixed-methods AI-supported intervention; chatbot feedback on peer-review comments	Improved comment quality and feedback literacy development
van den Bos & Tan (2019)	Dutch second-year university students	Quasi-experimental; anonymous vs. non-anonymous online peer review	Anonymity increased higher-order feedback and improved final module grades
Tran & Ma (2024)	Vietnamese university EFL students	Qualitative/mixed classroom study; self-regulated learning (SRL)-based online peer feedback training	Fostered engagement, strategy use, and peer help-seeking
Campbell & Batista (2023)	South Korean undergraduate EFL students	Controlled online peer-editing intervention with coded feedback	Supported social relatedness and self-efficacy, but raised accuracy concerns

The 19 core studies were concentrated in higher education EFL or L2 English writing settings and combined peer feedback with training, guidance, scaffolding, criteria support, technology support, or revision-linked structures. Their designs included quasi-experimental, controlled, mixed-methods, qualitative, and comparative studies. Because the studies differed substantially in research contexts, writing tasks, intervention designs, and outcome measures, the evidence base was methodologically heterogeneous. Therefore, this review adopted narrative thematic synthesis rather than meta-analysis to integrate the literature.

3.4 Data Extraction and Synthesis

Data were extracted on author, publication year, research context, participants, writing task, peer feedback design, type of training or scaffolding, technological platform, research design, reported outcomes, key conclusions, and limitations. Guided by the three research questions, the 19 core studies were coded and synthesized thematically (Braun & Clarke, 2006). The codes were organized into three dimensions: instructional design features, reported learning outcomes, and pedagogical challenges or implementation conditions. Frequency counts were used only to describe the occurrence of each category and were not used to estimate statistical effect sizes.

4. Results

4.1 RQ1: Instructional Designs of Peer Feedback Training

The synthesis identified seven instructional design features that appeared repeatedly across the 19 core studies. Table 6 presents descriptive frequency counts for each feature. Because a single study could contain multiple features, the percentages do not sum to 100%. The coding results indicate that peer

feedback training was commonly designed as a staged scaffolding process involving criteria or rubric support, teacher modeling, guided peer review, peer dialogue, technology- or AI-supported assistance, and revision-reflection linkage.

Table 6: Instructional design features of peer feedback training

Design feature and frequency	Representative core studies	Synthesis interpretation
Criteria/rubric/checklist support (n = 5/19; 26.3%)	Deng & Sitthitikul (2025); Hanh & Tinh (2022); Liu & Luo (2025); Tran & Ma (2024); Weng et al. (2024)	Made writing standards visible and helped students organize comments.
Teacher modeling/teacher support (n = 5/19; 26.3%)	Sun & Wang (2025); Min (2016); Cui et al. (2021); Zhang & Yu (2022); Campbell & Batista (2023)	Helped students calibrate feedback quality and increased confidence in peer review.
Guided peer-review sheets/prompts (n = 4/19; 21.1%)	Deng & Sitthitikul (2025); Hanh & Tinh (2022); Guo et al. (2024); Tran & Ma (2024)	Shifted peer feedback from general praise to more specific and actionable comments.
Peer dialogue and discussion (n = 5/19; 26.3%)	Deng & Sitthitikul (2025); Liu & Luo (2025); Sun & Wang (2025); Hassani et al. (2023); Zhang & Yu (2022)	Supported clarification, negotiation of meaning, and revision decisions.
Technology-supported peer feedback (n = 7/19; 36.8%)	Liu & Luo (2025); Zheng (2025); Awada & Diab (2023); Hassani et al. (2023); van den Bos & Tan (2019); Tran & Ma (2024); Campbell & Batista (2023)	Improved traceability and supported synchronous/asynchronous peer interaction.
AI-supported scaffolding (n = 3/19; 15.8%)	Lin & Hwang (2025); Guo et al. (2024); Guo et al. (2025)	Provided procedural support and feedback-on-feedback during peer review.
Revision-reflection linkage (n = 5/19; 26.3%)	Sun & Wang (2025); Zheng (2025); Awada & Diab (2023); Hassani et al. (2023); Tran & Ma (2024)	Connected peer comments with uptake, text revision, and self-regulated writing.

Table 6 shows that technology-supported peer feedback was the most frequently reported design feature. However, technology itself was not the core pedagogical mechanism. Digital tools appeared to be useful only when they were combined with rubrics, teacher modeling, guided prompts, peer dialogue, and revision-reflection tasks. Therefore, peer feedback training should be understood as a scaffolded instructional process rather than a single peer assessment activity.

4.2 RQ2: Reported Learning Outcomes

The outcome indicators reported in the included studies were not limited to writing scores. Table 7 summarizes five outcome categories: writing performance, feedback quality, revision uptake or revision quality, feedback literacy capacities, and affective or metacognitive outcomes. Because several studies reported more than one type of outcome, the categories are not mutually exclusive.

Table 7: Reported learning outcomes across the core studies

Outcome category and frequency	Representative core studies	Overall pattern	Caution
Writing performance (n = 8/19; 42.1%)	Deng & Sitthitikul (2025); Lin & Hwang (2025); Ghaneiarani et al. (2024); Hanh & Tinh (2022); Awada & Diab (2023); Cui et al. (2021); Guo et al. (2024); Hassani et al. (2023)	Mostly positive or mixed-positive outcomes were reported when peer feedback was trained and linked to revision.	Measures varied across argumentative writing, academic writing, paragraph writing, and writing accuracy.
Feedback quality (n = 6/19; 31.6%)	Do (2024); Min (2016); Guo et al. (2024); Zhang & Yu (2022); Guo et al. (2025); Tran & Ma (2024)	Training, modeling, SFL-informed support, rubrics, and AI prompts supported more specific and useful feedback.	Some studies assessed peer-review skill rather than final writing performance.
Revision uptake/revision quality (n = 6/19; 31.6%)	Sun & Wang (2025); Zheng (2025); Awada & Diab (2023); Hassani et al. (2023); van den Bos & Tan (2019); Tran & Ma (2024)	Feedback value depended on whether learners understood, trusted, and acted on peer comments.	Revision uptake indicators were not reported consistently; future studies should distinguish between comment adoption rate, successful revision rate, revision quality, and justified non-adoption of peer comments.
Feedback literacy capacities (n = 4/19; 21.1%)	Liu & Luo (2025); Weng et al. (2024); Guo et al. (2025); Tran & Ma (2024)	OPF and AI-supported review appeared to support criteria understanding, evaluative judgment, and feedback use.	Some studies did not explicitly use the term feedback literacy.

Outcome category and frequency	Representative core studies	Overall pattern	Caution
Motivation, self-efficacy, engagement (n = 4/19; 21.1%)	Zheng (2025); Cui et al. (2021); Tran & Ma (2024); Campbell & Batista (2023)	Peer feedback training may support affective and metacognitive outcomes such as motivation, self-efficacy, and engagement.	Many affective outcomes were based on self-report.

Overall, Table 7 shows that writing performance and feedback quality were the most frequently reported outcome categories. Feedback literacy and affective outcomes were reported less often, but they appeared increasingly in recent studies. Because the included studies varied in writing tasks, scoring rubrics, outcome measures, and research designs, the evidence should be interpreted as supportive rather than conclusive. Standardized reporting on feedback adoption – including comment adoption rate, successful revision rate, revision quality, and learners' justified acceptance or rejection of peer comments – would improve comparability across future studies.

4.3 RQ3: Pedagogical Challenges and Implementation Conditions

The synthesis also shows that the effectiveness of peer feedback training depends strongly on classroom implementation conditions. Table 8 summarizes six recurring challenges: learner readiness, feedback quality, revision uptake, classroom feasibility, technology integration, and AI-supported scaffolding. These findings indicate that successful peer feedback instruction requires not only sound instructional design but also feasible and reusable classroom routines.

Table 8: Challenges and implementation conditions

Challenge category and scope	Specific issue	Pedagogical implication
Learner readiness (cross-study theme)	Limited criteria understanding that low confidence, reluctance to criticize, low trust in peers.	Use criteria familiarization, sample comments, teacher modeling, and feedback language support.
Feedback quality (cross-study theme)	Vague praise, surface-level correction, inaccurate or unsupported suggestions.	Use rubrics, guided prompts, and revision-oriented comment frames.
Revision uptake (cross-study theme)	Students may ignore, misunderstand, or selectively use peer comments.	Require revision plans, reflection sheets, or feedback-use logs.
Classroom feasibility (cross-study theme)	Large classes, time pressure, uneven participation, teacher workload.	Adopt lightweight cycles, stable peer groups, and sample-based teacher monitoring.
Technology integration (cross-study theme)	Tools may record feedback but do not guarantee quality.	Use technology only when it supports recording, dialogue, revision, and monitoring.

Challenge category and scope	Specific issue	Pedagogical implication
AI-supported scaffolding (3 AI-related studies; 15.8%)	AI can scaffold comments but may not fully evaluate context, accuracy, or rhetorical intention.	Treat AI as support for peer/teacher interaction rather than as a replacement.

These implementation challenges suggest that classroom feasibility should be treated as a core design consideration rather than a secondary administrative issue. In large or time-constrained writing classes, peer feedback training can be organized as a streamlined, cyclical process embedded in regular writing instruction. Stable peer groups, simplified rubrics, sample-based teacher monitoring, and revision logs can reduce teacher workload while making students' use of peer suggestions more visible.

5. Discussion

5.1 How Peer Feedback Training Supports L2 English Writing

The synthesized findings suggest that peer feedback training supports L2 English writing through five mechanisms. First, clear evaluation criteria help students understand writing standards and reduce overly subjective peer comments. Second, structured feedback prompts make student comments more specific and revision oriented. Third, peer dialogue allows writers and reviewers to clarify meaning, discuss revision options, and connect comments to textual problems. Fourth, feedback language scaffolds help students provide constructive suggestions while reducing face-threatening remarks. Fifth, revision-reflection tasks connect peer feedback with text revision and feedback uptake (Winstone et al., 2017).

5.2 Comparison with Previous Studies and Reviews

The conclusions drawn in this review are consistent with previous studies that found that L2 writing instruction, systematic guidance, clear evaluation criteria, and structured revision opportunities can enhance the effectiveness of peer feedback. This finding aligns with Yu and Lee's (2016) review of peer feedback in second language writing and with Min's (2006, 2016) empirical work showing that peer review training can improve students' revision skills and evaluative abilities. This conclusion also corresponds to the student feedback literacy framework, which emphasizes that learners are not passive recipients of comments but need to recognize the value of feedback, make evaluative judgments, manage affective responses, and act on feedback in revision (Carless & Boud, 2018; Winstone et al., 2017).

The review further indicates that recent technology- and AI-supported studies have restructured peer feedback training into a traceable, scaffolded, and revision-linked process. This finding is consistent with reviews of technology-enhanced English language education, which suggest that digital tools have pedagogical value only when integrated with task design, teacher guidance, and student participation (Leong et al., 2024; Ramalingam et al., 2022; Tan & Yunus, 2023; Yingbao et al., 2025). Across the included studies, digital tools primarily

supported comment recording, peer communication, and transparency in the revision process.

However, these findings should be interpreted cautiously. In some studies, improvements in feedback quality, feedback literacy, motivation, or self-efficacy were more evident than improvements in final writing scores. Several studies also relied heavily on self-reported data rather than standardized writing measures. These differences indicate that structured peer feedback instruction should not be assumed to be automatically effective; its effectiveness depends on the quality of instructional support, students' trust in peer comments, teacher monitoring, time allocation, and the connection between feedback and revision. The main academic value of this review lies in integrating instructional design, learning mechanisms, and classroom implementation conditions into a framework for university-level L2 English writing research.

5.3 Pedagogical Implications

From a pedagogical perspective, peer feedback should be treated not as a casual exchange of opinions but as a systematically designed instructional practice. Teachers remain central as task designers, feedback models, and process monitors. Rubrics and benchmarking criteria clarify writing standards, teacher modeling shows students what counts as constructive and context-appropriate feedback, and revision-reflection tasks help students translate peer suggestions into concrete text revisions.

Although digital platforms and AI can support progress tracking, procedural scaffolding, and students' analytical skills, they cannot replace teacher guidance or student-to-student interaction. Therefore, the framework proposed in this review prioritizes lightweight classroom workflows, teacher monitoring, and revision-oriented feedback activities. It does not assume that automated tools alone can ensure effective peer collaboration.

5.4 Practical Recommendations for Classroom Implementation

At the classroom level, peer feedback instruction can be designed as a streamlined iterative process that includes a brief rubric explanation, teacher modeling of sample comments, scaffolded peer review, short peer discussion, and revision planning or feedback-use logs. In large-class settings, sample-based teacher monitoring is more feasible than reviewing every student comment one by one. Teachers can review selected comments, identify common problems, and provide whole-class feedback on specificity, politeness, accuracy, and revision usefulness.

AI-supported tools should be used with clear boundaries. If students delegate peer-reviewing responsibilities to an AI chatbot, the social interaction, peer trust, and critical thinking that peer feedback is intended to develop may be weakened. Teachers should thus position AI as support for planning, checking, or improving feedback language, not as a substitute reviewer. Students may submit the original peer comment, any AI-supported revision, and a brief explanation of how the final comment was evaluated and used.

5.5 Proposed Lightweight Peer Feedback Training Framework

Based on the mechanisms and pedagogical implications discussed above, this review proposes a lightweight peer feedback training framework for higher education L2 English writing. The framework is intended to help teachers integrate structured peer feedback instruction into regular writing lessons without adding excessive workload. Figure 2 presents the overall structure of the framework.

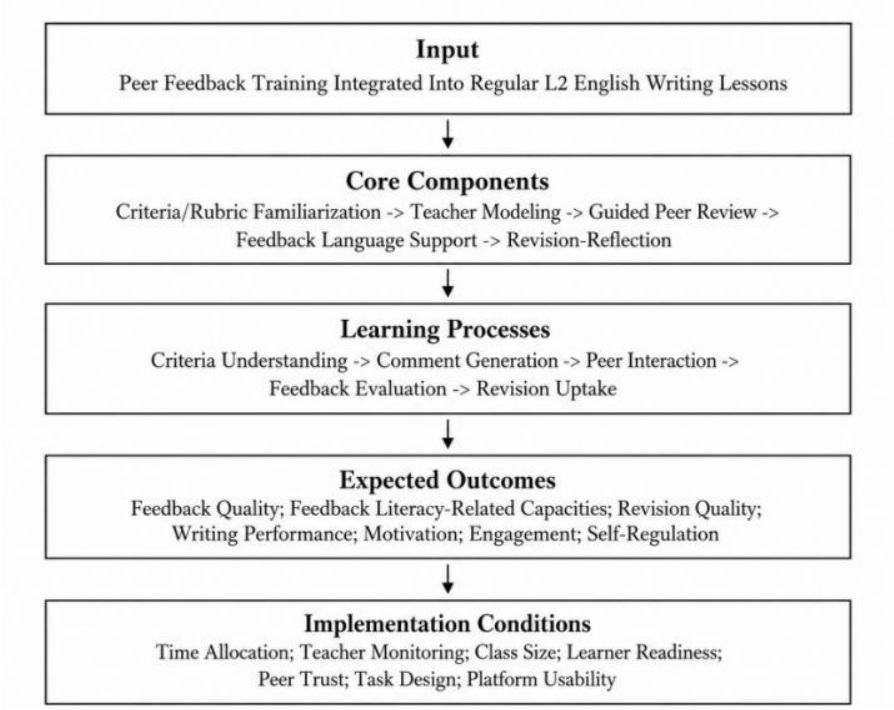

Figure 2: Lightweight peer feedback training framework for L2 English writing

Figure 2 presents the framework as a pedagogically synthesized model rather than a tested causal model. Table 9 illustrates how the framework can be implemented without replacing core writing instruction.

Table 9: Suggested classroom implementation cycle for peer feedback training

Phase	Suggested time	Classroom purpose
Criteria/rubric familiarization	5–8 minutes	Make writing standards visible before students review peer drafts.
Teacher modeling of sample comments	8–10 minutes	Show the difference between vague praise, surface correction, and revision-oriented feedback.
Guided peer review	15–20 minutes	Use a checklist or feedback sheet to help students generate specific comments.
Peer dialogue and clarification	8–10 minutes	Allow writers and reviewers to ask questions and negotiate revision decisions.
Revision planning or feedback-use log	5–8 minutes	Require students to state which comments they will adopt, adapt, or reject and why.
Teacher sample monitoring	5 minutes in class or after-class sampling	Reduce workload by reviewing selected comments and giving whole-class feedback.

5.6 Directions for Future Research

Future research should empirically pilot the proposed lightweight framework in higher education L2/EFL writing classrooms through quasi-experimental, longitudinal, or design-based studies. Researchers should measure not only writing scores but also feedback quality, feedback literacy, revision uptake, and learners' justified acceptance or rejection of peer comments. More studies are also needed in large-class EFL settings, especially with non-English majors and lower-proficiency learners. To improve comparability, future studies should report training length, teacher modeling procedures, grouping strategies, feedback sheet design, AI-use boundaries, revision requirements, comment adoption rate, successful revision rate, and revision quality.

6. Conclusion

This review synthesized 19 empirical studies on peer feedback training in higher education L2 English writing from 2016 to 2025. Regarding RQ1, peer feedback training was implemented through varied instructional designs, including criteria or rubric familiarization, teacher modeling, guided review sheets, peer dialogue, technology-supported review, AI-supported scaffolding, and revision-reflection activities. Regarding RQ2, the studies reported positive or mixed-positive outcomes for feedback quality, revision uptake, feedback literacy capacities, motivation, self-efficacy, engagement, and writing performance, although the strength of evidence varied across research designs and outcome measures. Regarding RQ3, effectiveness depended on learner readiness, peer trust, feedback quality, training duration, teacher monitoring, and feasible classroom routines.

The main contribution of this review is its training-focused synthesis and the proposed lightweight framework, which translates the thematic findings into a practical design for higher education EFL and L2 writing contexts. The framework may also help classroom teachers adapt peer feedback training into feasible, repeatable routines. Future intervention studies can examine the framework empirically and use more standardized indicators to measure feedback uptake, feedback literacy, and writing development. This analytical framework is theoretically grounded, but it has not yet been empirically tested as a causal model.

7. Limitations

This review has several limitations. Database coverage was limited to Scopus, ERIC, and Web of Science, and only English language publications were included. The included studies were also heterogeneous in terms of research design, intervention duration, writing tasks, outcome measures, and analytical procedures; therefore, the findings should be interpreted as thematic patterns rather than pooled causal evidence. In addition, some full-text reports could not be retrieved, which may have introduced selection bias. To reduce this risk, the review reported the search and screening process transparently and limited the core synthesis only to studies for which full-text evidence could be examined.

Conflicts of Interest
The authors declare no conflict of interest.

8. Acknowledgements

The authors acknowledge the use of QuillBot for language editing and improving the clarity of the manuscript. This paper remains an accurate representation of the authors' work and intellectual contributions.

9. References

Awada, G. M., & Diab, N. M. (2023). Effect of online peer review versus face-to-face peer review on argumentative writing achievement of EFL learners. *Computer Assisted Language Learning*, *36*(1–2), 238–256. https://doi.org/10.1080/09588221.2021.1912104

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Campbell, C. W., & Batista, B. (2023). To peer or not to peer: A controlled peer-editing intervention measuring writing self-efficacy in South Korean higher education. *International Journal of Educational Research Open, 4*, 100218. https://doi.org/10.1016/j.ijedro.2022.100218

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Cui, Y., Schunn, C. D., Gai, X., Jiang, Y., & Wang, Z. (2021). Effects of trained peer vs. teacher feedback on EFL students' writing performance, self-efficacy, and internalization of motivation. *Frontiers in Psychology, 12*, 788474. https://doi.org/10.3389/fpsyg.2021.788474

Deng, Y., & Sitthitikul, P. (2025). The effects of the guided dialogic peer feedback-based writing instruction on Chinese EFL students' writing performance in an integrated blended learning environment. *rEFLections, 32*(1), 1–27. https://doi.org/10.61508/refl.v32i1.277804

Do, H. M. (2024). E-peer feedback training on L2 writing: A case study of inexperienced learners. *The JALT CALL Journal*, *20*(1), 1–20. https://doi.org/10.29140/jaltcall.v20n1.1242

Ghaneiarani, S., Alavi, S. M., & Kaivanpanah, S. (2024). Enhancing writing ability of Iranian EFL learners through learning-oriented assessment: Peer and teacher feedback implementation. *Language Testing in Asia, 14*, Article 39. https://doi.org/10.1186/s40468-024-00298-w

Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education, 63*, 100962. https://doi.org/10.1016/j.iheduc.2024.100962

Guo, K., Zhang, E. D., Li, D., & Yu, S. (2025). Using AI-supported peer review to enhance feedback literacy: An investigation of students' revision of feedback on peers' essays. *British Journal of Educational Technology, 56*(4), 1612–1639. https://doi.org/10.1111/bjet.13540

Hanh, L. T. T., & Tinh, B. T. (2022). Applying peer-review checklist to improve Vietnamese EFL university students' writing skills. International Journal of Learning, *Teaching and Educational Research, 21*(5), 166–181. https://doi.org/10.26803/ijlter.21.5.9

Hassani, J. B., Chalak, A., & Heidari Tabrizi, H. (2023). Investigating the effect of synchronous and asynchronous computer-mediated peer feedback on IELTS candidates' writing development and interactions. *Education and Self Development, 18*(1), 26–42. https://doi.org/10.26907/esd.18.1.03

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101. https://doi.org/10.1017/S0261444806003399

Leong, L. V., Yunus, M. M., & Ismail, H. H. (2024). Integration of techno-pedagogical approach in English as a second language classroom: A systematic review. *International Journal of Evaluation and Research in Education, 13*(6), 4394–4408. https://doi.org/10.11591/ijere.v13i6.29976

Lin, C.-J., & Hwang, G.-J. (2025). Artificial intelligence-supported procedural scaffolding for promoting EFL learners' writing performance in flipped peer assessment activities. *Interactive Learning Environments,* 1–15. https://doi.org/10.1080/10494820.2025.2532629

Liu, F., & Luo, W. (2025). Exploring the effects of online peer feedback on tertiary students' peer feedback literacy in English writing: Focusing on design elements. *Studies in Educational Evaluation, 87*, 101516. https://doi.org/10.1016/j.stueduc.2025.101516

Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118–141. https://doi.org/10.1016/j.jslw.2006.01.003

Min, H.-T. (2016). Effect of teacher modeling and feedback on EFL students' peer review skills in peer review training. *Journal of Second Language Writing, 31*, 43–57. https://doi.org/10.1016/j.jslw.2016.01.004

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Ramalingam, S., Yunus, M. M., & Hashim, H. (2022). Blended learning strategies for sustainable English as a second language education: A systematic review. *Sustainability, 14*(13), 8051. https://doi.org/10.3390/su14138051

Sun, H., & Wang, M. (2025). Effects of teacher intervention and type of peer feedback on student writing revision. *Language Teaching Research, 29*(3), 1094–1113. https://doi.org/10.1177/13621688221080507

Tan, S. Y., & Yunus, M. M. (2023). Sustaining English language education with social networking sites (SNSs): A systematic review. *Sustainability, 15*(7), 5710. https://doi.org/10.3390/su15075710

Tran, T. T. T., & Ma, Q. (2024). Online peer feedback training based on self-regulated learning in English as a foreign language writing: Perceived usefulness and students' engagement. *Studies in Educational Evaluation, 83*, 101418. https://doi.org/10.1016/j.stueduc.2024.101418

Van den Bos, A. H., & Tan, E. (2019). Effects of anonymity on online peer review in second-language writing. *Computers & Education, 142,* 103638. https://doi.org/10.1016/j.compedu.2019.103638

Weng, F., Zhao, C. G., & Chen, S. (2024). Effects of peer feedback in English writing classes on EFL students' writing feedback literacy. *Assessing Writing, 61*, 100874. https://doi.org/10.1016/j.asw.2024.100874

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

Yingbao, L., Ismail, H. H., & Yunus, M. M. (2025). Teaching pedagogies using ICT in university academic English writing across the globe – A systematic literature

review. *International Journal of Learning, Teaching and Educational Research*, *24*(3), 656–678. https://doi.org/10.26803/ijlter.24.3.31

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching, 49*(4), 461–493. https://doi.org/10.1017/S0261444816000161

Zhang, X., & Yu, S. (2022). Training student writers in conducting peer feedback in L2 writing: A meaning-making perspective. *Applied Linguistics Review, 13*(4), 439–460. https://doi.org/10.1515/applirev-2019-0045

Zheng, Q. (2025). Promoting draft revision motivation and quality through technology-enhanced peer feedback workshops. *Educational Technology Research and Development, 73*, 1225–1243. https://doi.org/10.1007/s11423-025-10450-9