Mediation Without a Mind: A Sociocultural Reconsideration of Student Feedback Literacy for Generative AI Feedback in ELT Writing

Ng Yong Sheng, Nur Syafiqah Yaccob

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.100601289

Received: 01 July 2026; Accepted: 06 July 2026; Published: 17 July 2026

ABSTRACT

Generative artificial intelligence increasingly provides immediate feedback to English language learners, yet learners' capacity to interpret, evaluate, and act on such feedback, known as feedback literacy, remains undertheorised in relation to this new source. This conceptual paper adopts a theory adaptation approach, using Carless and Boud's four-dimensional model of student feedback literacy as the domain theory and Vygotsky's sociocultural theory as the method theory. It examines how appreciating feedback, making judgements, managing affect, and taking action are reshaped when feedback is generated by general-purpose artificial intelligence rather than by teachers or peers. The paper argues that the model remains useful, but that each dimension faces additional strain because generative artificial intelligence functions as a mediating artefact and cannot be assumed to perform the pedagogical role of a more knowledgeable other. Although it can provide responsive assistance, it does not reliably identify a learner's zone of proximal development, interpret developmental needs, or provide support that is carefully adjusted and gradually withdrawn. Learners must therefore undertake more of the evaluative, emotional, and regulatory work required to use feedback effectively. In English language writing contexts, target-language proficiency influences how strongly these limitations are experienced, creating a proficiency paradox in which learners who depend most on artificial intelligence feedback may have fewer linguistic resources to judge its accuracy, relevance, and appropriateness. The paper uses illustrative ELT writing scenarios and indicative proficiency benchmarks to anchor this argument, suggesting that the paradox may be most acute for learners at CEFR A2-B1, while recognising that such boundaries are gradual rather than fixed. It concludes that feedback literacy for generative artificial intelligence should be deliberately developed through teacher-supported mediation, guided verification, reflective revision, comparison of human and artificial intelligence feedback, and selective adaptation of suggestions.

Keywords: English language teaching, feedback literacy, generative AI, sociocultural theory, zone of proximal development

INTRODUCTION

Some English language learners increasingly use generative artificial intelligence (GenAI) tools to support writing-related activities (Dizon et al., 2025). Recent empirical studies indicate that English language learners can use GenAI tools such as ChatGPT and Gemini, as well as artificial intelligence (AI) or AI-enhanced writing platforms such as Grammarly, not only to generate or develop written texts but also to obtain rapid feedback on their drafts (Mahapatra, 2024; Tran, 2025; Mekheimer, 2025). For learners who once waited days for an intermittent teacher response, the appeal of feedback that is instant, private, and apparently inexhaustible is obvious. Although research has increasingly examined the affordances, quality, and learning effects of GenAI-generated feedback, comparatively less is known about learners' capacity to interpret, evaluate, and use such feedback in their subsequent work (Zhan & Yan, 2025; Zou et al., 2024). This concern is visible in learners' actual feedback practices. In a study of 16 undergraduates who used ChatGPT feedback to revise IELTS writing tasks, Zhan and Yan (2025) recorded 308 feedback prompts, but only five, or 1.6%, concerned the verification of ChatGPT's credibility. Most interactions were also limited to a single round, leading the authors to characterise learners' behavioural engagement with the feedback as superficial. These findings demonstrate that

access to immediate GenAI feedback does not necessarily result in sustained verification or critical engagement, highlighting a concrete need to develop learners' feedback literacy in ELT writing.

Research on feedback has consistently indicated that feedback quality alone does not guarantee learning; its effectiveness depends substantially on learners' capacity to make sense of feedback and act on it (Carless & Boud, 2018; Hyland & Hyland, 2006). The understandings, capacities, and dispositions required to engage productively with feedback are commonly conceptualised as feedback literacy (Carless & Boud, 2018). Emerging research suggests that this capacity is also central to engagement with GenAI-generated feedback: learners differ in how they seek, evaluate, modify, and apply GenAI feedback, and stronger feedback literacy is associated with more productive feedback uptake and perceptions of feedback usefulness (Hawkins et al., 2025; Mendoza et al., 2026; Yang et al., 2025; Zhan et al., 2025; Zhan & Yan, 2025).

However, Carless and Boud's (2018) influential framework predates the widespread adoption of conversational GenAI and was developed within a pre-GenAI higher-education feedback context, with its practical illustrations centred mainly on peer feedback, engagement with exemplars, teacher facilitation, and learner self-evaluation. Although the framework acknowledges that feedback may reach learners from automated sources, the automated feedback available at the time consisted largely of deterministic automated writing evaluation and grammar-checking tools that flagged a bounded range of features against fixed rules. Conversational, general purpose GenAI differs in kind: it generates novel, fluent responses on demand via probabilistic next-token prediction and may appear confident even when it contains inaccurate or fabricated information (Farquhar et al., 2024; Kalai et al., 2026). Applying the framework to this newer form of feedback therefore requires reconsideration, as it places additional evaluative demands on learners, who must judge not only the feedback's meaning and usefulness but also its credibility and accuracy. Such demands may be especially consequential in English language teaching (ELT) writing, where learners must evaluate feedback communicated in a target language in which their proficiency is still developing.

This paper argues that Carless and Boud's (2018) four-component framework, including appreciating feedback, making judgements, managing affect, and taking action, remains a productive analytical lens, but that the operation of each component is reshaped and, in some respects, strained when feedback is generated by general-purpose GenAI. The paper further argues that these strains share a common root, which becomes visible when GenAI feedback is examined through the lens of sociocultural theory. From this perspective, GenAI feedback can serve as a mediating artefact, but it should not be treated automatically as a more knowledgeable other (MKO). This paper does not attempt to resolve the philosophical question of whether GenAI genuinely understands the language it produces. Instead, it examines whether current general-purpose systems can perform the pedagogical functions expected of an MKO in feedback interactions. Although GenAI can produce fluent, responsive, and apparently personalised feedback, this does not establish that it can reliably interpret a learner's intended meaning, diagnose developmental needs, identify the learner's zone of proximal development, or adjust and gradually withdraw assistance as the learner develops (Hicks et al., 2024; Vendrell & Johnston, 2026; Yang, 2026). This limitation has direct implications for ELT. Learners may mistake fluent and personalised output for dependable pedagogical judgement and may therefore accept feedback that is inaccurate, inappropriate, or poorly matched to their proficiency. Teachers consequently remain important for verifying GenAI-generated feedback, interpreting learner needs, and providing contingent guidance that supports increasing learner independence and self-regulation (Choi, 2025; Ding et al., 2025).

This limitation may weaken an important process through which external support contributes to internalisation. In the ELT context, it is compounded by a proficiency paradox: learners with lower target-language proficiency may face greater cognitive constraints when evaluating feedback delivered in that language (Wu & Zhao, 2025), even though they may need such feedback most. Target-language proficiency thus functions as a condition that intensifies the consequences of the mediator's limitations, a relationship that is developed later in the paper. To keep the argument concrete, the discussion is anchored in three illustrative ELT writing scenarios and, where proficiency is concerned, in indicative benchmark levels drawn from the Common European Framework of Reference for Languages (Council of Europe, 2020), with approximate IELTS equivalents. The aim here is not to dismantle the framework but to extend its application to the distinctive conditions of GenAI-mediated feedback in ELT writing.

Research Questions

This paper examines how student feedback literacy changes when feedback in ELT writing is generated by artificial intelligence rather than provided by a teacher or peer. It addresses the following research questions:

1. How is student feedback literacy reshaped when the source of feedback is generative AI rather than a human in ELT writing?

2. What are the implications of these changes for ELT writing pedagogy?

The paper adopts a conceptual rather than empirical approach. It follows the theory-adaptation design described by Jaakkola (2020), in which an established theory is examined and extended through a second theoretical perspective. Following the distinction drawn by Lukka and Vinnari (2014), Carless and Boud's (2018) student feedback literacy framework serves as the domain theory, and Vygotsky's (1978) sociocultural theory serves as the method theory through which the framework is reconsidered.

LITERATURE REVIEW

Generative AI Feedback in ELT Writing

Feedback is central to the teaching of second-language writing, but providing timely, individualised, and pedagogically effective teacher feedback can be difficult, particularly in large classes (Hyland & Hyland, 2006; Mahapatra, 2024). Conversational GenAI appears to loosen some of these constraints by providing immediate, accessible, and on-demand feedback, while reducing the social anxiety or fear of judgement that may accompany feedback-seeking from teachers or peers (Zhan & Yan, 2025). Empirical work has begun to document both its promise and its hazards. GenAI-assisted feedback has been associated with more frequent revision and gains in organisation and cohesion (Mekheimer, 2025), stronger writing self-efficacy (Huang & Mizumoto, 2024), broader improvements in academic writing and positive learner perceptions (Mahapatra, 2024), and reduced writing anxiety (Wang, D., 2024). The same literature, however, raises concerns about over-reliance, passive or uncritical acceptance of suggestions, weakened evaluative agency, and possible threats to learner autonomy (Hamamra et al., 2026; Mekheimer, 2025; Zhan & Yan, 2025; Zulkefli & Ismail, 2025).

Mahapatra (2024) additionally identifies concerns about the loss of creativity, the imposition of writing patterns, and academic integrity. Taken together, these findings suggest that the same features that make GenAI feedback attractive- its immediacy, persuasive presentation, and continuous availability may also make over-dependence more plausible. These systems can provide corrective feedback on local features such as grammar, lexis, and mechanics and, to varying degrees, commentary on organisation, cohesion, and style. In conversational systems such as ChatGPT, the feedback is also shaped by learners' prompts, meaning that it develops through interaction rather than functioning solely as a fixed judgement.

The Emerging Intersection of Feedback Literacy and AI

A small but growing body of research has connected feedback literacy with generative AI. Zhan and Yan (2025) found that students employed more cognitive than metacognitive strategies when engaging with ChatGPT feedback. Their trust in the feedback varied, and their behavioural engagement in prompting, interacting with ChatGPT, and revising their writing was often superficial. The authors argue that feedback literacy in a GenAI context should encompass prompt engineering, evaluative judgement, emotional reflexivity, ethical decision-making, and metacognitive skills. Zhan et al. (2025) further propose that GenAI's capacity to enhance feedback engagement depends on the alignment between the opportunities and constraints of the GenAI feedback environment and learners' feedback literacy.

The present paper develops this discussion specifically in relation to L2 writing and examines each of Carless and Boud's (2018) four dimensions of student feedback literacy separately. More importantly, it offers a sociocultural account of why GenAI feedback places additional demands on L2 learners. It further theorises a shared source underlying these demands and the resulting proficiency paradox. While existing research has identified these competencies and related them to the affordances and constraints of GenAI feedback, the present

paper seeks to explain, from a sociocultural perspective, why they become necessary in L2 writing and how they may be shaped by learners' target-language proficiency.

Student Feedback Literacy

Nieminen and Carless (2023) describe feedback literacy research as concerned with the capacities of students and teachers to optimise the benefits of feedback opportunities, while cautioning against treating these capacities as purely individual and decontextualised attributes. A particularly influential conceptualisation is provided by Carless and Boud (2018), who define student feedback literacy as the understandings, capacities, and dispositions needed to make sense of information and use it to enhance work or learning strategies. Crucially, the concept shifts agency towards the learner, recasting the learner from a passive recipient to an active participant who must undertake cognitive, affective, and behavioural work for feedback to contribute meaningfully to learning.

Carless and Boud (2018) specify four interrelated features: appreciating feedback, which involves understanding its value, recognising one's active role, and acknowledging that feedback takes different forms and comes from multiple sources; making judgements, which involves developing the evaluative judgement needed to assess one's own and others' work; managing affect, which involves regulating the emotions provoked by feedback rather than responding defensively; and taking action, through which feedback information is translated into improvement.

Subsequent research has empirically extended this model (Molloy et al., 2020) and proposed a framework for the teacher capacities needed to create conditions that support student feedback literacy (Carless & Winstone, 2023). Although Carless and Boud acknowledge the social construction of feedback understanding and identify dialogue, peer interaction, exemplars, modelling, and coaching as developmental resources, their framework does not elaborate in detail the sociocultural mechanisms through which external mediation becomes internalised as learner regulation. This omission matters because feedback literacy is associated with identifiable self-regulatory processes rather than exposure to feedback alone. In a study of 1,975 university learners across 34 Chinese institutions, Chen et al. (2026a) found that self-evaluation predicted appreciation of feedback and evaluative judgement, whereas goal setting and task strategies predicted feedback uptake. Without explaining how such processes are socially supported and progressively internalised, the framework describes the capacities associated with feedback literacy more clearly than it explains how those capacities develop into independent learner regulation. Sociocultural theory therefore provides a complementary lens through which these developmental processes can be examined.

Sociocultural Theory

Sociocultural theory (SCT) originated in Vygotsky's cultural-historical psychology and was subsequently extended and systematised for second-language research, notably by Lantolf and Thorne (2006). This paper draws particularly on four central SCT concepts: mediation, the more knowledgeable other, the zone of proximal development, and internalisation. Mediation refers to the principle that human cognition and activity are organised through culturally developed tools and signs, especially language (Lantolf & Xi, 2023; Poehner & Leontjev, 2023). These mediational resources do not merely facilitate learners' contact with their environment; through their use, learners may reorganise how they understand, plan, and regulate activity (Leontjev & deBoer, 2025). The zone of proximal development (ZPD) refers to the distance between what learners can accomplish independently and what they can achieve through adult guidance or collaboration with more capable peers (Vygotsky, 1978; Poehner & Lantolf, 2021). The more knowledgeable other (MKO) is the person, typically a teacher, expert, or more capable peer, whose guidance within this zone makes such assisted performance possible; crucially, the MKO does not merely supply answers but interprets the learner's responses and adjusts assistance to their developing understanding. Appropriately mediated activity within the ZPD can initiate developmental processes that later become available for independent performance (Poehner & Lantolf, 2021; Zhang, 2023). Internalisation is the transformative process through which socially mediated activity is reconstructed on the individual plane, enabling control to shift progressively from other-regulation towards self-regulation (Lantolf & Thorne, 2006).

This framework is applicable to feedback in L2 writing. In a foundational empirical study, Aljaafreh and Lantolf (1994) conceptualised corrective feedback as graduated, contingent, and dialogically negotiated mediation within the learner's ZPD. The expert seeks to provide the minimum necessary assistance, increases explicitness only when necessary, and adjusts support based on the learner's responsiveness. As learners gain control, the relevance of explicit assistance diminishes and responsibility for performance shifts towards the learner. From this perspective, effective feedback is not defined primarily by the quantity of information supplied but by whether assistance is calibrated to the learner's developmental needs and promotes increasing self-regulation. Han and Li (2024) provide a partial empirical illustration of teacher-mediated GenAI feedback. In their study, teachers reviewed and adapted ChatGPT-generated output, converted direct corrections into coded indirect feedback, and supported learners in successfully revising a range of local and rhetorical problems. However, because the study examined learners' immediate revisions across two writing tasks rather than the gradual withdrawal of assistance or transfer to later independent writing, it did not establish whether the support was internalised as independent competence (Han & Li, 2024).

Applied to GenAI-generated feedback, sociocultural theory offers a useful lens for distinguishing assistance that promotes learner development and increases self-regulation from assistance that completes part of the task while leaving regulatory control with the external tool. In this paper, student feedback literacy provides the substantive framework identifying the capacities learners need to engage with feedback, while sociocultural theory provides an explanatory and analytical framework for examining how mediated assistance may or may not contribute to learning, internalisation, and greater self-regulation. Examining the four dimensions of student feedback literacy through this perspective illuminates both the potential and the limitations of GenAI-generated feedback in ELT writing.

Reconsidering The Four Dimensions

Three Illustrative Scenarios

Before each dimension is examined, the paper's central concern is made concrete through three short scenarios. These are constructed illustrations rather than reported cases, but each is consistent with behaviour documented in the empirical literature, including GenAI marking acceptable language as erroneous, offering corrections that are unnecessary or beyond learners' proficiency levels, and reformulating text in ways that alter meaning or style (Cengiz & Aptoula, 2026; ElEbyary & Shabara, 2024; Lin & Crosthwaite, 2024). The scenarios are referred to throughout the discussion that follows, and they are offered as plausible instances rather than as claims about how any particular system behaves in every case.

Scenario 1: the misread idiom. A learner at approximately CEFR B1 writes, in a narrative essay, "Even now, my grandmother's advice still rings in my ears whenever I face a difficult decision". The GenAI tool flags the idiom as unclear and proposes "Even now, I still remember my grandmother's advice whenever I face a difficult decision". The original sentence is grammatically and idiomatically acceptable, and the proposed revision, while fluent, removes a rhetorical choice the learner made deliberately after encountering the idiom in class. Because the learner's command of English idiom is still developing, they have limited independent means of determining whether the flagged expression was actually erroneous, and the most economical response is to accept the change. The text becomes slightly blander, and an opportunity to confirm the correct use of an idiom is lost. Nothing in the exchange is dramatic, which is precisely the difficulty: the miscalibration is plausible, minor, and hard for this learner to detect.

Scenario 2: the smoothed voice. A Malaysian undergraduate writing a reflective essay describes her childhood: "In my kampung, the river was our swimming pool, our playground and our fridge". Asked to improve the paragraph, the tool returns: "In my village, the river served several purposes, from recreation to a source of food". The revision is grammatically impeccable and rhetorically flat: the local term, the three-part rhythm, and the small surprise of "fridge" carry the writer's voice, and all three are smoothed away. This pattern is consistent with findings that GenAI-assisted revision can move learner writing towards formulaic, homogenised prose and weaken cognitive ownership (Dung, 2025; Mai, 2026). A learner who cannot yet articulate why the original was better may accept the revision and, over repeated cycles, may learn to write towards the tool's preferences rather than towards her own developing voice.

Scenario 3: assistance above the learner's level. A learner at approximately CEFR A2 pastes a paragraph into a conversational GenAI tool and asks for feedback. The response praises the draft and then recommends "foregrounding the counterargument in a concessive subordinate clause before the refutation", supplying a fully rewritten paragraph pitched at roughly C1 level. The metalinguistic explanation is beyond the learner's current comprehension, and the rewritten paragraph contains structures the learner could neither have produced nor evaluated independently. The learner therefore does what the interaction makes easiest and replaces the original paragraph wholesale. Feedback delivered above learners' levels is documented empirically (Cengiz & Aptoula, 2026), and wholesale replacement of this kind resembles the superficial behavioural engagement observed by Zhan and Yan (2025). The submitted text improves; it is far less clear that the learner does.

The surface problems differ across the three scenarios, involving an unnecessary correction, the erosion of voice, and level-inappropriate assistance respectively, but their underlying structure is shared. In each case the mediating artefact produces fluent, confident output that is not calibrated to the individual learner, and the learner lacks either the linguistic resources or the evaluative criteria to notice what has gone wrong. The four dimensions of feedback literacy are reconsidered below with these scenarios in view.

Appreciating Feedback

In Carless and Boud's (2018) framework, appreciating feedback involves recognising its value and understanding one's active role in making sense of and using it. Their broader conception acknowledges that feedback information may come from multiple sources, including automated systems, though the automated feedback prominent when the framework was developed was largely rule-based and deterministic. Conversational GenAI raises an additional question concerning the nature and limitations of its source. Human feedback can draw on a reader's situated interpretation of a text's communicative purpose, context, and intended meaning, whereas GenAI output is generated through probabilistic language modelling that is not inherently oriented towards truth or grounded in a situated understanding of the writer's purpose (Hicks et al., 2024). Responsibility for GenAI-generated judgements is also less straightforwardly located in the system itself and instead extends across the human and organisational agents involved in its development and use (Constantinescu & Kaptein, 2025). Appreciating GenAI feedback appropriately, therefore, involves recognising its genuine affordances while remaining alert to its limitations and treating it as a potentially useful but fallible source rather than an unquestionable authority. Scenario 1 illustrates why this stance matters in practice: feedback that is confidently and fluently phrased may still misjudge what constitutes acceptable language, so appreciating GenAI feedback requires recognizing its fallibility.

In ELT settings, this evaluative demand may be difficult for learners who lack sufficient linguistic knowledge or appropriate guidance. Mun (2024) found that learners believed inadequate grammatical and lexical knowledge could make it difficult to determine whether ChatGPT's suggestions were accurate, while Zhan and Yan (2025) identified varying levels of trust and emphasised the need to develop evaluative judgement and metacognitive engagement. From a sociocultural perspective, this paper therefore proposes that GenAI-generated feedback should first be understood as a mediating artefact and should not be automatically assumed to function as an MKO in a manner equivalent to that of a human teacher. Raine (2025) argues that GPT-based systems may act as MKOs to a significant extent, but only when their feedback is appropriately calibrated to learners' current levels and learners actively engage with it. Ding et al. (2025) and Choi (2025) further indicate that GenAI feedback remains limited in its understanding of authorial intent, contextual nuance, emotional meaning, and higher-order content, and that these limitations can make teacher involvement or hybrid feedback particularly valuable.

Making Judgements

This is arguably the dimension that GenAI transforms most substantially and one in which the particular demands of the ELT context become especially visible. With teacher feedback, learners may still need to assess relevance, clarity, and usefulness, but perceived linguistic and instructional competence can provide an initial basis for trust. Papi et al. (2026) found that competence-based teacher trust increased the perceived value of written corrective feedback and promoted feedback monitoring and inquiry, while Zeevy-Solovey (2024) found that learners commonly preferred teacher feedback because they regarded the teacher as an expert and a reliable source.

Teacher feedback may therefore carry a stronger initial presumption of credibility, although learners must still interpret it critically and decide how to act on it.

GenAI-generated feedback complicates this presumption. ChatGPT can mark acceptable language as erroneous, provide incorrect or level-inappropriate corrections, vary its feedback approach, and generate unnecessary or redundant suggestions (Cengiz & Aptoula, 2026; ElEbyary & Shabara, 2024; Lin & Crosthwaite, 2024). Scenario 1 is an instance of precisely this behaviour, in which an acceptable idiom is flagged as an error and revised away. Learners have also expressed concerns about its reliability and the risk of overreliance (Mun, 2024). Because GenAI information may appear plausible and relevant even when unreliable, making judgements must include comparing feedback with other sources and evaluating the validity of the feedback itself (Zhan & Yan, 2025).

In ELT contexts, this requirement creates a proficiency-related tension. Evaluating feedback about a target-language feature requires sufficient linguistic knowledge to recognise whether a proposed correction is plausible, yet learners seek feedback partly because that knowledge is still developing. Yan and Zhang's (2024) multiple-case study found that higher-proficiency learners used metacognitive monitoring and evaluation more effectively, whereas lower-proficiency learners had greater difficulty integrating these strategies into feedback processing and revision. In that study, technological competence appeared to partly compensate for limited linguistic proficiency through more intensive interaction with ChatGPT, suggesting that proficiency is important but does not independently determine feedback engagement (Yan & Zhang, 2024). Evidence from automated writing evaluation similarly indicates that learners' responses to and perceptions of feedback vary according to language proficiency (Xu & Zhang, 2022). This is consequential because GenAI feedback is not consistently reliable or developmentally appropriate: Cengiz and Aptoula (2026) found that 6.35% of ChatGPT feedback instances were incorrect, while a further 5.54% were unnecessary or beyond the learners' proficiency levels. Taken together, these findings suggest that learners with more limited linguistic resources may face particular difficulty evaluating plausible but inaccurate or level-inappropriate suggestions, although direct proficiency-based comparisons of error detection remain needed.

To give this proficiency-related tension more concrete boundaries, it can be expressed in terms of the Common European Framework of Reference for Languages (Council of Europe, 2020), provided any such mapping is read as indicative rather than absolute. CEFR global descriptors characterise basic users (A1–A2) as handling short, simple and largely formulaic language; independent users as producing simple connected text on familiar topics at B1 and clear, detailed text with more flexible language use at B2; and proficient users (C1–C2) as exercising fluent, well-structured control of the language, including nuances of register (Council of Europe, 2020). Verifying GenAI feedback typically requires a learner to compare a proposed form against an internal model of the target language and, often, to parse a metalinguistic explanation delivered in English. On this basis, the proficiency paradox is likely to be most acute for learners in the approximate A2 to B1 range, often loosely aligned with IELTS overall bands of roughly 3.5 to 5.0. Such learners are typically proficient enough to solicit and partially comprehend GenAI feedback on extended writing, yet may lack the grammatical and lexical resources to distinguish accurate corrections from plausible but inaccurate or unnecessary ones, as in Scenarios 1 and 3. Learners around B2, roughly IELTS 5.5 to 6.5, may verify feedback on frequent structures with increasing success while remaining vulnerable in areas such as register, idiomaticity, and discourse organisation, whereas learners at C1 and above, roughly IELTS 7.0 or higher, generally retain more of the resources needed to treat GenAI output as one voice among several. These bands should be read as gradient tendencies rather than thresholds: no single level marks the point at which evaluative capacity appears or disappears, and proficiency interacts with feedback literacy, metacognitive strategy use, and technological competence (Xiangming & Wei, 2026; Yan & Zhang, 2024). The claim advanced here is therefore conditional rather than categorical. The wider the gap between the learner's proficiency and the linguistic level of the text and feedback under evaluation, the more likely the required evaluative work will exceed the resources available for it, and the greater the need for external corroboration.

Sociocultural theory further clarifies this tension. In Aljaafreh and Lantolf's (1994) account, learners' capacity to regulate their performance and evaluate possible corrections is not treated as something that must already be fully developed before receiving feedback. It can develop through graduated and contingent assistance, as a more knowledgeable other adjusts support in response to learner performance and progressively transfers

responsibility to the learner. General-purpose GenAI may simulate some outward features of this process by providing immediate, individualised responses and allowing learners to seek clarification or refine feedback through successive prompts (Slamet, 2024; Yan & Zhang, 2024). However, current evidence does not establish that such systems can independently diagnose a learner's developmental level, interpret their needs pedagogically, or fade assistance in response to demonstrated development (Fang & Han, 2025). In Godoy's (2026) intervention, for example, responsibility was transferred from GenAI support to learners through teacher-designed offline revision rather than through ChatGPT's autonomous withdrawal. The system's apparent adaptation may therefore partly reflect the information and instructions supplied through prompting rather than a pedagogically grounded understanding of the learner's readiness.

Taken together, these constraints are twofold. ELT learners may have limited linguistic resources for evaluating GenAI feedback, while GenAI cannot be assumed to provide the calibrated mediation through which stronger evaluative judgement might develop. In this sense, GenAI can function as a mediating artefact without consistently fulfilling the pedagogical role of a more knowledgeable other. For lower-proficiency learners, the appropriate disposition is therefore not necessarily confident independent judgement but calibrated caution. This involves recognising when they lack sufficient knowledge or confidence to evaluate a particular piece of GenAI feedback and seeking corroboration from teachers, reliable reference materials, or other authoritative sources. Recognising the limits of one's current ability to judge feedback is itself an important form of evaluative judgement.

Managing Affect

In Carless and Boud's (2018) framework, managing affect involves maintaining emotional equilibrium and avoiding defensiveness when receiving critical feedback. GenAI may initially appear favourable in this respect because its immediate and accessible interaction can be experienced by some learners as a low-stakes, non-judgemental form of support, although such affective responses are not uniformly positive (Du & Chen, 2026; Chen et al., 2026b; Teng, 2025). Removing some interpersonal threats, however, may introduce new risks. One is over-trust: minimal emotional resistance and easy access to repeated suggestions may coincide with superficial processing or acceptance without sufficient scrutiny (Zhan & Yan, 2025; Du & Chen, 2026). A second concern is voice. When ELT writers rely heavily on GenAI-generated phrasing, immediate gains in fluency and linguistic accuracy may come at the cost of originality, individual expression, and cognitive ownership (Mai, 2026). Scenario 2 illustrates the dynamic: each individual revision appears reasonable, yet the cumulative effect may be a text that sounds progressively less like its writer. Related evidence suggests that AI-assisted writing may increase confidence and reduce stress, while also raising concerns about inaccurate output, fabricated information, and the potential loss of academic voice (Ravi & Mohamad, 2026).

Sociocultural theory helps explain this risk. Internalisation involves actively reconstructing socially mediated resources at the individual level, integrating them with existing knowledge, beliefs, and experiences to make them personally meaningful (Shabani & Bakhoda, 2024). From this perspective, wholesale acceptance of GenAI-generated phrasing may inhibit internalisation when language is incorporated without comprehension, critical evaluation, or meaningful adaptation, a concern consistent with Mai's (2026) findings on formulaic writing, reduced cognitive effort, and weakened ownership. Managing affect in a GenAI context, therefore, also involves regulating trust: neither defensive rejection nor uncritical deference, but a calibrated stance that the convenience and psychological reassurance of GenAI can sometimes make difficult to sustain.

Taking Action

With human feedback, obstacles to action can include difficulty interpreting less explicit comments, recognising their relevance, and translating them into appropriate revisions (Kaya, 2023; Ma, 2023). GenAI can reduce some of the procedural effort involved by supplying corrections, reformulations, and alternative passages on demand. Although these affordances can facilitate revision, they may also reduce the cognitive work of interpreting feedback, generating alternatives, and deciding how to revise. Such work is potentially productive because effective feedback engagement involves noticing and understanding feedback, monitoring one's response, and making revision decisions rather than merely enacting supplied corrections (Cheng & Xu, 2025; Zhang & Hyland, 2023). When GenAI-generated language is adopted without evaluation or meaningful adaptation, uptake may

remain superficial, with limited monitoring and reflection (Zhan & Yan, 2025) and a greater risk of dependency and weakened evaluative control (Hamamra et al., 2026). Action may then resemble transcription more than revision. Scenario 3 is the limiting case, in which the learner's sole action is the wholesale replacement of the original paragraph, an act that changes the text without exercising any of the interpretive or evaluative processes that make action developmentally productive.

From a sociocultural perspective, mediated L2 development can involve a shift from other-regulation towards self-regulation as externally provided assistance is internalised (Aljaafreh & Lantolf, 1994; Lantolf & Thorne, 2006). When learners adopt GenAI-generated language with little evaluation or reconstruction, the interaction may instead sustain other-regulation by transferring substantial control over wording and revision to the tool. Learners may follow GenAI-generated forms closely or make only cursory changes rather than independently making and justifying revision decisions (Zhan & Yan, 2025; Hamamra et al., 2026). As noted in relation to managing affect, textual performance may consequently improve without corresponding development in linguistic understanding, critical engagement, or cognitive ownership (Mai, 2026). The concern about overreliance is therefore not simply that learners use GenAI feedback, but that they may act on it in ways that weaken the movement towards self-regulated performance.

Taking action effectively, therefore, requires deliberate, selective uptake that keeps learners engaged in the regulation process. Lukešová and Jennings (2026) designed a clue-based strategy in which learners infer corrections from guided hints rather than passively receive direct answers, thereby sustaining active problem-solving and reflective engagement. Similarly, Zhang and Liu (2026) recommend self-evaluation before GenAI consultation, comparison of learner- and GenAI-generated alternatives, and graduated prompting from implicit to more explicit assistance. Prompts that request clues, explanations, or relevant criteria rather than replacement text can turn feedback uptake into an opportunity for reflection and possible internalisation. In this way, the artefact may support movement towards self-regulation rather than substitute for the learner's regulatory activity.

The Common Root

Two connected observations underlie the strains discussed above, and it is worth being explicit about how they relate. The more fundamental of the two is a sociocultural limit on GenAI-mediated assistance. General-purpose systems generate language through probabilistic modelling rather than through human-like conceptual understanding, and should not be automatically equated with a human MKO (Vendrell & Johnston, 2026). Although they can reproduce some outward features of adaptive support, current evidence does not establish that they can independently infer a learner's ZPD or systematically adjust and fade assistance in response to demonstrated development. Even in structured domains, the accuracy and helpfulness of personalised feedback vary with task complexity, error type, and the contextual information provided to the system (Reddig et al., 2025). This limitation is the common root of the four strains: because the mediator does not reliably interpret the learner and calibrate support, the evaluative, affective, and regulatory work that a human MKO would ordinarily share falls more heavily on the learner. Teachers therefore remain important for interpreting learner needs and helping students understand the capabilities and limitations of GenAI-generated assistance (Kilde-Westberg et al., 2025).

Target-language proficiency is best understood not as a second, independent root but as the condition that determines how exposed a given learner is to this limitation. Because the learner must now supply more of the interpretive and evaluative work themselves, the linguistic resources available to them become decisive. Proficiency can influence how learners process feedback, regulate cognitive and metacognitive activity, formulate feedback requests, and translate suggestions into revisions (Xu & Zhang, 2022; Yan & Zhang, 2024). It may also shape some positive emotional responses to GenAI feedback, although feedback literacy appears to be a more consistent affective predictor than writing proficiency itself (Xiangming & Wei, 2026), and technological competence may partly compensate for limited proficiency (Yan & Zhang, 2024). Yuan and Dao (2026) further show that relatively proficient learners may engage in both selective evaluation and complete uptake, although their study did not compare proficiency levels. Proficiency is therefore an important but not the only factor shaping learners' engagement with GenAI feedback; it interacts with feedback literacy and technological competence rather than operating independently. Feedback literacy for GenAI in L2 contexts should thus not be understood simply as a general construct with an additional technological component; it is

also conditioned by learners' developing command of the target language. In the indicative benchmark terms introduced earlier, exposure to the mediator's limitations is likely to be greatest where the gap between the demands of verification and the learner's available resources is widest, tentatively below B2, and to narrow, without disappearing, at higher levels.

This limitation in GenAI-mediated assistance places additional interpretive, evaluative, affective, and regulatory demands on learners across all four dimensions of student feedback literacy (Zhan et al., 2025; Zhan & Yan, 2025). Appreciation requires learners to recognise the nature and limits of the mediator; judgement requires them to verify assistance that may not be developmentally calibrated; managing affect requires them to regulate trust and maintain ownership; and taking action requires them to remain actively involved in monitoring and revision. GenAI-generated feedback has genuine value, but that value depends partly on learners' capacity to evaluate, monitor, and regulate its use (Zhan & Yan, 2025).

Implications For ELT Writing Pedagogy

It follows that feedback literacy for GenAI-generated feedback should not be assumed to develop automatically and may be especially difficult to develop without support among lower-proficiency learners. The apparent convenience and immediacy of generative AI can obscure the evaluative and metacognitive work required to use its feedback productively (Wang et al., 2025; Zhan & Yan, 2025). Lower-proficiency learners may have fewer linguistic resources to interpret and critically evaluate automated suggestions (Anastasia et al., 2024; Lu et al., 2025; Yan & Zhang, 2024). Taken together, these findings suggest that feedback literacy should be developed deliberately through contingent mediation rather than assumed to emerge solely from access to a general-purpose GenAI tool.

This gives the teacher's role in developing students' feedback literacy a more specific form. Carless and Winstone (2023) conceptualise this role as the design of feedback environments and the provision of guidance, coaching, and modelling. In a GenAI-mediated context, this includes pedagogical guidance that the artefact itself may not reliably provide (Carless & Winstone, 2023; Han & Li, 2024). Making judgements, arguably the dimension under greatest strain, therefore requires structured and scaffolded intervention rather than general encouragement alone (Bearman et al., 2024; Zhang & Liu, 2026). Drawing on these principles, learners could compare teacher and GenAI feedback on the same text, identify where the two sources agree or differ, and discuss the standards and reasoning underlying those differences. As a further pedagogical extension, teachers could provide GenAI-generated commentary containing deliberate errors for learners to identify and explain.

Through such teacher-structured activities, learners may develop and calibrate their evaluative judgement by comparing feedback sources and considering the standards underlying their decisions (Bearman et al., 2024; Wood & Pitt, 2025). Drawing on contingent-teaching principles, teachers can initially provide assistance responsive to learners' current understanding and progressively reduce this support as learners develop greater evaluative and self-regulatory capacity (Behzadi Soufiani et al., 2025). For taking action, feedback uptake should similarly be made visible and reflective rather than automatic. A revision log that records which GenAI suggestions learners accepted, adapted, or rejected, along with their reasons, may support more reflective and selective revision (Sharshova et al., 2025). Such an approach is consistent with recommendations for process-oriented AI-supported writing assessment that makes learners' reasoning and decision-making visible while retaining human judgement over complex aspects of writing (Singh et al., 2026).

The indicative benchmarks proposed earlier also suggest a gradation of support, offered here as a heuristic rather than a prescription. For learners at approximately A2 to B1, teachers might pre-review and adapt GenAI feedback before learners encounter it, as in Han and Li's (2024) teacher-mediated design, and might confine verification tasks to features already within learners' developing repertoire. For learners around B2, structured comparison of teacher and GenAI feedback on the same text, together with the planted-error tasks described above, may offer an appropriate level of challenge. For learners at C1 and above, revision logs and independent corroboration with reliable reference sources often suffice, with teacher mediation focused on register, voice, and disciplinary conventions. In every case, the underlying aim is the same: to keep the evaluative work within, rather than beyond, the learner's zone of proximal development, and to withdraw support contingently as capacity develops (Behzadi Soufiani et al., 2025; Zhang & Liu, 2026).

Encouraging learners to adapt GenAI suggestions to their intended meaning and voice, rather than reproduce GenAI-generated wording directly, may promote more critical engagement and help preserve their authorial voice and agency (Dung, 2025; Wang, 2025). Because feedback literacy encompasses both dispositions, such as valuing feedback and regulating trust and emotion, and practical capacities for interpreting and using feedback, appreciating feedback and managing affect may be fostered through explicit, teacher-guided discussion (Gozali et al., 2024; Zhan & Yan, 2025). Explicit discussion of how generative AI produces text, why fluent and persuasive output may still require verification, and why apparent linguistic competence does not guarantee reliable understanding may help learners develop a more accurate appreciation of GenAI feedback and avoid uncritical trust (Lund et al., 2025; Zhan & Yan, 2025). Discussion of voice, agency, and ownership may also help learners consider whether GenAI-generated language has been understood, adapted, and meaningfully integrated into their own developing linguistic resources (Dung, 2025; Wang, 2025).

Across these practices, the teacher's role changes but does not diminish. Teachers can exercise contextual pedagogical judgement by reviewing and adapting GenAI-generated feedback to learners' needs, while structuring opportunities for learners to develop the feedback literacy needed to engage productively with GenAI as a mediating resource (Han & Li, 2024; Thiruchelvan & Zakaria, 2026; Zhan et al., 2025). Evidence from Malaysian ESL teachers further indicates that although AI-powered writing assistants may support revision and learner independence, concerns about inaccurate feedback, over-reliance, insufficient training, and unequal infrastructure reinforce the need for continued teacher mediation (Thiruchelvan & Zakaria, 2026).

CONCLUSION

This paper asked how student feedback literacy must be reconceived when feedback in ELT writing is generated by a probabilistic machine rather than provided by a teacher or peer. Taking Carless and Boud's (2018) four-dimensional model as its domain theory and Vygotsky's (1978) sociocultural theory as its method theory, it argued that the model remains productive but that each dimension is reshaped, and in places strained, once feedback issues from generative AI are considered. Appreciating feedback must now recognise the tool as a mediating artefact rather than grant it the standing of a more knowledgeable other; making judgements expands to include evaluating the validity of the feedback itself; managing affect comes to involve regulating trust and protecting voice; and taking action must stay deliberate if uptake is not to become transcription that holds the learner in permanent other-regulation.

Beneath these strains lies a common root, with target-language proficiency as the condition that governs how sharply it is felt. Most fundamentally, GenAI mediates without a mind in the functional sense at issue here: it generates language probabilistically, without the learner's situated knowledge, and thus cannot reliably infer a zone of proximal development or calibrate and withdraw assistance as a human expert can. Because the mediator cannot reliably supply this calibrated support, more of the evaluative and regulatory work falls to the learner, and target-language proficiency then shapes how well that work can be done across all four dimensions. Making judgements is therefore arguably the dimension under greatest strain, meeting a proficiency paradox that sociocultural theory deepens, since learners seek feedback precisely because their command of the language is incomplete, yet evaluating that feedback presupposes the knowledge they lack.

The practical conclusion is that feedback literacy of this kind cannot be left to emerge on its own, least of all among lower-proficiency learners, who may rely most on GenAI feedback yet be least equipped to evaluate it. It must be cultivated deliberately, through the responsive and contingent mediation that the tool itself cannot guarantee. This repositions rather than diminishes the teacher, who remains the more knowledgeable other in the feedback relationship; not necessarily by replacing the tool, but by providing the interpretive and contingent guidance that GenAI can support but not itself supply, frequently working in combination with it. Comparing teacher and GenAI feedback on the same text, locating planted errors in GenAI-generated commentary, and logging which suggestions were accepted, adapted, or rejected and why are concrete means of building evaluative judgement and keeping learners within the regulation of their own writing.

The argument offered here is conceptual, which is both its limitation and its point. It provides a reasoned reinterpretation rather than evidence of how ELT learners in fact engage with GenAI feedback. The expanded burden of judgement, the proficiency paradox, the threat to appropriation, and the risk of sustained other-

regulation are each hypotheses that classroom research can test. So too are the indicative proficiency bands proposed here: whether the paradox is indeed most acute in the approximate A2 to B1 range, and how far feedback literacy and technological competence moderate it, are empirical questions that the benchmarks are intended to make testable rather than to settle. As generative AI becomes a permanent feature of the ELT writing environment, understanding how learners can be helped to use it well, and not merely to use it, is a task the field can no longer postpone.

REFERENCES

1. Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465–483. https://doi.org/10.2307/328585
2. Anastasia, W., Murtisari, E., & Isharyanti, N. (2024). Lower-proficiency EFL students' use of Grammarly in writing: Behavior, cognition, and affect. *The JALT CALL Journal*, 20(1), 1–25. https://doi.org/10.29140/jaltcall.v20n1.1089
3. Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321
4. Behzadi Soufiani, L., Ahangari, S., & Saeidi, M. (2025). The effect of the model of contingent teaching on improving Iranian EFL learners' essay writing. *Language Teaching Research*, 29(5), 2161–2176. https://doi.org/10.1177/13621688221113021
5. 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
6. Carless, D., & Winstone, N. (2023). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 28(1), 150–163. https://doi.org/10.1080/13562517.2020.1782372
7. Cengiz, Y., & Aptoula, N. Y. (2026). Comparing AI and human feedback at higher education: Level appropriateness, quality and coverage. *Language Teaching*, 59(1), 122–127. https://doi.org/10.1017/S0261444825000199
8. Chen, S., Zhang, L., & Li, A. W. (2026a). Leveraging self-regulated learning to support university students' feedback literacy in blended learning environments. *Assessment & Evaluation in Higher Education*, 51(1), 174-193. https://doi.org/10.1080/02602938.2025.2546349
9. Chen, Z., Wei, W., & Zou, D. (2026b). Generative AI technology and language learning: Global language learners' responses to ChatGPT videos in social media. *Interactive Learning Environments*, 34(2), 907–920. https://doi.org/10.1080/10494820.2025.2511248
10. Cheng, X., & Xu, J. (2025). Engaging second language (L2) students with synchronous written corrective feedback in technology-enhanced learning contexts: A mixed-methods study. *Humanities and Social Sciences Communications*, 12(1), 712. https://doi.org/10.1057/s41599-025-05007-3
11. Choi, W. (2025). Exploring high school students' preferences for English writing feedback: Non-native English teacher vs. native English teacher vs. ChatGPT. *Sage Open*, 15(4), 21582440251383693. https://doi.org/10.1177/21582440251383693
12. Constantinescu, M., & Kaptein, M. (2025). Responsibility gaps, LLMs & organisations: Many agents, many levels, and many interactions. *Science and Engineering Ethics*, 31(1), 36. https://doi.org/10.1007/s11948-025-00560-1
13. Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. https://www.coe.int/lang-cefr
14. Ding, L., Zou, D., & Kohnke, L. (2025). ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing. *Education and Information Technologies*, 30, 26777–26799. https://doi.org/10.1007/s10639-025-13775-3
15. Dizon, G., Gold, J., & Barnes, R. (2025). ChatGPT for self-regulated language learning: University English as a foreign language students' practices and perceptions. *Digital Applied Linguistics*, 3, 102510. https://doi.org/10.29140/dal.v3.102510

16. Du, J. L., & Chen, Y. (2026). Exploring English for academic purposes learners' engagement with ChatGPT-4 in academic writing revision: A case study from a private language institute in China. *Frontiers in Artificial Intelligence*, 9, 1802007. https://doi.org/10.3389/frai.2026.1802007
17. Dung, L. Q. (2025). Syntactic and rhetorical transformation in Vietnamese EFL students' academic writing. *Journal of Education and E-Learning Research*, 12(4), 628–635. https://doi.org/10.20448/jeelr.v12i4.7848
18. ElEbyary, K., & Shabara, R. (2024). ChatGPT-generated corrective feedback: Does it do what it says on the tin? *Teaching English with Technology*, 24(3), 68–89. https://doi.org/10.56297/vaca6841//BFFO7057/MYEH4562
19. Fang, S., & Han, Z. (2025). On the nascency of ChatGPT in foreign language teaching and learning. *Annual Review of Applied Linguistics*, 45, 148–178. https://doi.org/10.1017/S026719052510010X
20. Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, 630(8017), 625–630. https://doi.org/10.1038/s41586-024-07421-0
21. Godoy, D. (2026). AI mediated scaffolding for second language academic writing in EFL composition classrooms. *Discover Education*, 5, 475. https://doi.org/10.1007/s44217-026-01386-0
22. Gozali, I., Wijaya, A. R. T., Lie, A., Cahyono, B. Y., & Suryati, N. (2024). Leveraging the potential of ChatGPT as an automated writing evaluation (AWE) tool: Students' feedback literacy development and AWE tools integration framework. *The JALT CALL Journal*, 20(1), 1–22. https://doi.org/10.29140/jaltcall.v20n1.1200
23. Hamamra, B., Khlaif, Z. N., & Mahamid, N. (2026). AI literacy and mediated learner autonomy in ChatGPT-supported EFL writing: A qualitative study at a Palestinian university. *International Journal of Educational Technology in Higher Education*, 23(1), 19. https://doi.org/10.1186/s41239-026-00597-7
24. Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System*, 126, 103502. https://doi.org/10.1016/j.system.2024.103502
25. Hawkins, B., Taylor-Griffiths, D., & Lodge, J. M. (2025). Summarise, elaborate, try again: Exploring the effect of feedback literacy on AI-enhanced essay writing. *Assessment & Evaluation in Higher Education*, 1–13. https://doi.org/10.1080/02602938.2025.2492070
26. Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26(1), 38. https://doi.org/10.1007/s10676-024-09775-5
27. Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics*, 1, 102324. https://doi.org/10.29140/dal.v1.102324
28. Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101. https://doi.org/10.1017/S0261444806003399
29. Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1), 18–26. https://doi.org/10.1007/s13162-020-00161-0
30. Kalai, A. T., Nachum, O., Vempala, S. S., & Zhang, E. (2026). Evaluating large language models for accuracy incentivizes hallucinations. *Nature*, 1047–1051. https://doi.org/10.1038/s41586-026-10549-w
31. Kaya, F. (2023). Understanding written teacher feedback in L2 writing in higher education: Perceptions, emotions and practices of pre-service English language teachers. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 52(3), 819–833. https://doi.org/10.14812/cuefd.1245489
32. Kilde-Westberg, S., Johansson, A., & Enger, J. (2025). Generative AI as a lab partner: A case study. *Physical Review Physics Education Research*, 21(2), 020119. https://doi.org/10.1103/ggy1-3kjk
33. Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
34. Lantolf, J. P., & Xi, J. (2023). Digital language learning: A sociocultural theory perspective. *TESOL Quarterly*, 57(2), 702–715. https://doi.org/10.1002/tesq.3218
35. Leontjev, D., & deBoer, M. (2025). Teacher as creator: Orchestrating the learning environment to promote learner development. *Language Teaching Research*, 29(6), 2371–2392. https://doi.org/10.1177/13621688221117654
36. Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System*, 127, 103529. https://doi.org/10.1016/j.system.2024.103529
37. Lu, Q., Yao, Y., Xiao, L., Zhu, X., & Yin, H. (2025). Exploring the impact of a GenAI-supported feedback practice on student feedback literacy in L2 writing. *Computer Assisted Language Learning*, 1–30. https://doi.org/10.1080/09588221.2025.2605541

38. Lukešová, A., & Jennings, P. J. (2026). Clue before correction: ChatGPT-enhanced strategy for promoting autonomous and reflective language learning. *Innovation in Language Learning and Teaching*, 1–29. https://doi.org/10.1080/17501229.2025.2612532
39. Lukka, K., & Vinnari, E. (2014). Domain theory and method theory in management accounting research. *Accounting, Auditing & Accountability Journal*, 27(8), 1308–1338. https://doi.org/10.1108/AAAJ-03-2013-1265
40. Lund, B., Mannuru, N. R., Teel, Z. A., Lee, T. H., Ortega, N. J., Simmons, S., & Ward, E. (2025). Student perceptions of AI-assisted writing and academic integrity: Ethical concerns, academic misconduct, and use of generative AI in higher education. *AI in Education*, 1(1), 2. https://doi.org/10.3390/aieduc1010002
41. Ma, M. (2023). Exploring learning-oriented assessment in EAP writing classrooms: Teacher and student perspectives. *Language Testing in Asia*, 13(1), 33. https://doi.org/10.1186/s40468-023-00249-x
42. Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9. https://doi.org/10.1186/s40561-024-00295-9
43. Mai, H. T. (2026). Evaluating the impact of using ChatGPT on EFL students' argumentative skills. *Discover Education*, 5(1), 189. https://doi.org/10.1007/s44217-026-01153-1
44. Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4(1), 170. https://doi.org/10.1007/s44217-025-00602-7
45. Mendoza, N. B., Xiong, Y., & Yan, Z. (2026). Making sense of AI feedback: How students' feedback literacy moderates the link between ChatGPT acceptance and self-regulated learning. *Educational Psychology*, 1–23. https://doi.org/10.1080/01443410.2026.2641521
46. Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45(4), 527–540. https://doi.org/10.1080/02602938.2019.1667955
47. Mun, C. Y. (2024). EFL learners' English writing feedback and their perception of using ChatGPT. *Journal of English Teaching through Movies and Media*, 25(2), 26–39. https://doi.org/10.16875/stem.2024.25.2.26
48. Nieminen, J. H., & Carless, D. (2023). Feedback literacy: A critical review of an emerging concept. *Higher Education*, 85(6), 1381–1400. https://doi.org/10.1007/s10734-022-00895-9
49. Papi, M., Turner, J., Song, W., Soliman, H., Mahbodi, M., & Yarbrough, A. (2026). A good relationship is not enough: How teacher trust shapes feedback-seeking behaviors in second language writing. *Journal of Second Language Writing*, 73, 101318. https://doi.org/10.1016/j.jslw.2026.101318
50. Poehner, M. E., & Lantolf, J. P. (2021). The ZPD, second language learning, and the transposition~transformation dialectic. *Cultural-Historical Psychology*, 17(3), 31–41. https://doi.org/10.17759/chp.2021170306
51. Poehner, M. E., & Leontjev, D. (2023). Peer interaction, mediation, and a view of teachers as creators of learner L2 development. *International Journal of Applied Linguistics*, 33(1), 18–32. https://doi.org/10.1111/ijal.12444
52. Raine, P. (2025). Using GPT-3 as a more knowledgeable other to provide feedback on English language learners' spoken transcripts. *TESL-EJ*, 29(2). https://doi.org/10.55593/ej.29114a7
53. Ravi, D. M., & Mohamad, M. (2026). Exploring the Use of Artificial Intelligence in Journal Article Writing Among Postgraduate Students at Universiti Kebangsaan Malaysia (UKM). International Journal of Research and Innovation in Social Science (IJRISS), 10(1), 3645-3658. https://dx.doi.org/10.47772/IJRISS.2026.10100286
54. Reddig, J. M., Arora, A., & MacLellan, C. J. (2025). Generating in-context, personalized feedback for intelligent tutors with large language models. *International Journal of Artificial Intelligence in Education*, 35, 3459–3500. https://doi.org/10.1007/s40593-025-00505-6
55. Shabani, K., & Bakhoda, I. (2024). Internalization and externalization in a computerized L2 context from Vygotskian optique. *Language Teaching Research Quarterly*, 46, 174–198. https://doi.org/10.32038/ltrq.2024.46.13
56. Sharshova, R., Salkhanova, Z., Yelubayeva, P., Maral, A., Sholakhova, A., & Ainur, K. (2025). The use of AI writing tools in second language learning to enhance Kazakh IT students' academic writing skills. *Forum for Linguistic Studies*, 7(8), 267–281. https://doi.org/10.30564/fls.v7i8.10408

57. Singh, R. K. V., Hashim, H., Jamaludin, K. A., & Jamil, M. R. M. (2026). Assessing AI-supported English academic writing in higher education: A narrative review of challenges, emerging practices, and prospective strategies. Arab World English Journal (AWEJ) Special Issue on Artificial Intelligence, 3, 436-551. https://dx.doi.org/10.24093/awej/A3.28
58. Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence*, 4(1), 46. https://doi.org/10.1007/s44163-024-00143-2
59. Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, 60(1), e12811. https://doi.org/10.1111/ejed.12811
60. Thiruchelvan, T., & Zakaria, N. Y. K. (2026). ESL Teachers' Perceptions of AI-Powered Writing Assistants in Language Learning Learning. I*nternational Journal of Education, Psychology and Counselling (IJEPC)*, 11(62), 1282–1295. https://doi.org/10.35631/IJEPC.1162074
61. Tran, T. T. T. (2025). Enhancing EFL writing revision practices: The impact of AI- and teacher-generated feedback and their sequences. *Education Sciences*, 15(2), 232. https://doi.org/10.3390/educsci15020232
62. Vendrell, M., & Johnston, S. K. (2026). Scaffolding critical thinking with generative AI: Design principles for integrating large language models in higher education. *Computers and Education: Artificial Intelligence*, 100572. https://doi.org/10.1016/j.caeai.2026.100572
63. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
64. Wang, C. (2025). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning*, 30(3), 1825–1846. https://doi.org/10.1007/s10758-024-09744-3
65. Wang, D. (2024). Teacher- versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *International Review of Research in Open and Distributed Learning*, 25(3), 37–56. https://doi.org/10.19173/irrodl.v25i3.7646
66. Wang, F., Li, N., Cheung, A. C., & Wong, G. K. (2025). In GenAI we trust: An investigation of university students' reliance on and resistance to generative AI in language learning. International *Journal of Educational Technology in Higher Education*, 22(1), 59. https://doi.org/10.1186/s41239-025-00547-9
67. Wood, J., & Pitt, E. (2025). Empowering agency through learner-orchestrated self-generated feedback. *Assessment & Evaluation in Higher Education*, 50(1), 127–143. https://doi.org/10.1080/02602938.2024.2365856
68. Wu, Z., & Zhao, J. (2025). Exploring learners' engagement with GenAI-generated feedback in EFL writing: A Chinese middle school case study. *Computer Assisted Language Learning*, 1–26. https://doi.org/10.1080/09588221.2025.2600035
69. Xiangming, L., & Wei, W. (2026). Interplay effects of AI feedback and individual differences on learners' emotional responses. *Education and Information Technologies*. https://doi.org/10.1007/s10639-026-13937-x
70. Xu, J., & Zhang, S. (2022). Understanding AWE feedback and English writing of learners with different proficiency levels in an EFL classroom: A sociocultural perspective. *The Asia-Pacific Education Researcher*, 31(4), 357–367. https://doi.org/10.1007/s40299-021-00577-7
71. Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. *Humanities and Social Sciences Communications*, 11(1), 1–14. https://doi.org/10.1057/s41599-024-03543-y
72. Yang, K. (2026). Generative AI as a virtual more-knowledgeable other: A mixed-methods investigation of DeepSeek's impact on grammatical accuracy and technology acceptance in Chinese ESL writing. *Social Reports*, 1(1), 110–128. https://doi.org/10.66638/ep1sq081
73. Yang, Y., Hussin, S., & Hashim, H. (2025). Feedback Literacy and EFL Learner Engagement With ChatGPT Feedback: Predicting Feedback Uptake and Perceived Usefulness. *Theory & Practice in Language Studies (TPLS)*, 15(11), 3760-3771. https://doi.org/10.17507/tpls.1511.30
74. Yuan, S., & Dao, P. (2026). ESL tertiary learners' engagement in ChatGPT-assisted second language writing. *Language Awareness*, 1–24. https://doi.org/10.1080/09658416.2026.2652333
75. Zeevy-Solovey, O. (2024). Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: Students' perceptions and preferences. *Technology in Language Teaching & Learning*, 6(3), 1482. https://doi.org/10.29140/tltl.v6n3.1482

76. Zhan, Y., Boud, D., Dawson, P., & Yan, Z. (2025). Generative artificial intelligence as an enabler of student feedback engagement: A framework. *Higher Education Research & Development*, 44(5), 1289–1304. https://doi.org/10.1080/07294360.2025.2476513
77. Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1–14. https://doi.org/10.1080/02602938.2025.2471821
78. Zhang, Y. (2023). Promoting young EFL learners' listening potential: A model of mediation in the framework of dynamic assessment. *The Modern Language Journal*, 107(S1), 113–136. https://doi.org/10.1111/modl.12824
79. Zhang, Y., & Liu, Y. (2026). Designing ChatGPT-mediated feedback activities in EFL writing: A design-based study of the dialogic feedback triangle. *Assessment & Evaluation in Higher Education*, 51(1), 137–160. https://doi.org/10.1080/02602938.2025.2571846
80. Zhang, Z. V., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. *Assessing Writing*, 58, 100784. https://doi.org/10.1016/j.asw.2023.100784
81. Zou, S., Guo, K., Wang, J., & Liu, Y. (2024). Investigating students' uptake of teacher- and ChatGPT-generated feedback in EFL writing: A comparison study. *Computer Assisted Language Learning*, 1–30. https://doi.org/10.1080/09588221.2024.2447279
82. Zulkefli, N. H., & Ismail, H. H. (2025). The Impact of AI on ESL Writing Autonomy: A Systematic Review. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 10(8), e003555-e003555. https://doi.org/10.47405/mjssh.v10i8.3555