
| RESEARCH ARTICLE

Assistant or Ghostwriter? EAP Instructors' Views on the Opportunities and Risks of GenAI in L2 Writing

Mohamed AboulEla Abdelmageed¹ ✉ & Alena Shannaq²

¹ Centre for Preparatory Studies, Dhofar University, Salalah, Dhofar, Oman; <https://orcid.org/0000-0003-3165-0711>

² Centre for Preparatory Studies, Dhofar University, Salalah, Dhofar, Oman; <https://orcid.org/0009-0002-3836-1787>

Corresponding Author: Mohamed AboulEla Abdelmageed **E-mail:** mabdelmageed@du.edu.om

| ABSTRACT

Generative artificial intelligence (GenAI) systems are able to generate, revise and comment on second language (L2) student writing. However, less is known about where L2 writing teachers locate the opportunities and risks of such systems across the writing process. To that end, this study examined how English for academic purposes (EAP) writing instructors at two universities in Oman position GenAI within the L2 writing process. It focused on, first, the stages at which instructors locate GenAI's opportunities and risks, second, their views on GenAI as a source of feedback, and lastly their preparedness to manage academic integrity in AI-assisted L2 writing. A cross-sectional questionnaire combining Likert-scale items and open-ended prompts was completed by 51 EAP instructors, of whom 49 were retained for analysis. Closed items were analysed with descriptive statistics, Spearman correlations, McNemar tests and Mann–Whitney U tests. Open-ended responses were coded by writing stage through qualitative content analysis. The results showed that L2 writing instructors located GenAI's opportunities mainly in their own preparation of materials and model texts and in feedback. On the other hand, they highlighted its risks in students' drafting and in assessment. Revising and editing was contested, cited equally often as an opportunity and a risk. EAP instructors were more confident in using GenAI to prepare materials than in designing students' use of it at each writing stage. Just over half agreed that AI improves their feedback. This view was related to their overall attitude towards GenAI but not to their confidence as feedback providers. Most instructors felt responsible for checking students' GenAI use and confident in setting boundaries, but their confidence in maintaining academic integrity was unrelated to this sense of responsibility. More EAP instructors reported knowing how to use AI detectors than understanding how such detectors work. The study concludes that EAP instructors' readiness to integrate GenAI into L2 writing teaching and learning is uneven across the L2 writing process. Professional learning and development should therefore target the pedagogical design of student-facing AI use, the evaluation of AI feedback, AI detection literacy, and L2 writing assessment redesign.

| KEYWORDS

Generative AI (GenAI), English for academic purposes (EAP), L2 writing, writing feedback, academic integrity, TPACK

| ARTICLE INFORMATION

ACCEPTED: 01 September 2026

PUBLISHED: 26 October 2026

DOI: 10.32996/bjal.2025.6.5.1

1. Introduction

Since late 2022, generative artificial intelligence (GenAI) systems such as ChatGPT have been able to generate, revise and comment on extended text in seconds. For second language (L2) writers, the potential benefits are considerable, including support with ideas, vocabulary and grammar, model texts, and immediate feedback on drafts (Barrot, 2023; Su et al., 2023). The risks are equally clear. Students may outsource the writing that teachers need to assess. In addition, AI output can be inaccurate or stylistically inappropriate. Moreover, the tools used to detect AI-generated text have proved unreliable (Barrot, 2023; Weber-Wulff et al., 2023), especially for writing by non-native learners of English (Liang et al., 2023).

How these benefits and risks materialise in L2 writing depends largely on teachers, who decide whether, when and how students may use GenAI. Teachers not only largely design most L2 writing learning and assessment tasks, but they also provide the feedback that guides L2 learners' revisions (Kohnke et al., 2023). Despite teachers' central role, most research on GenAI in L2 writing has examined students' use of the tools or compared AI-generated feedback with teacher feedback (Crosthwaite & Sun,

2026; Steiss et al., 2024). Far less is known about how English for academic purposes (EAP) teachers themselves locate GenAI within the L2 writing process, that is, which stages they see it serving, where they see it threatening learning, and how prepared they feel to manage its consequences for academic integrity. Answers to these questions also indicate EAP writing teachers' professional learning and development (PLD) needs.

These questions are particularly significant in the Gulf region, including Oman. English is the medium of instruction in most Omani higher education institutions, and students who complete Arabic-medium secondary schooling enter general foundation programmes that prepare them for academic writing in English (Naghdi-pour, 2021). Omani students have encountered persistent difficulties with vocabulary, grammar, paraphrasing and summarising (Mustafa et al., 2022). Many turn to AI tools for translation, grammar checking and idea generation (Al-Raimi et al., 2024). As for Omani English teachers, recent studies have suggested that they prefer to use AI for preparing writing instructional materials rather than for feedback or grading (Khattak et al., 2025). Additionally, Omani EFL faculty have reported a strong sense of ethical responsibility despite a marked lack of institutional guidance (Vellanki & Tabassum, 2026). However, there is a considerable lack of regional studies on how EAP instructors view GenAI at each stage of L2 writing teaching and learning.

This study therefore addresses this gap through a survey of 49 English instructors at two universities in southern Oman. The primary research question is "How do EAP instructors view GenAI across the stages of L2 writing teaching and learning?" Three sub-questions follow the teaching and learning of L2 writing from the teacher preparation of instructional material, students' drafting and revising, and feedback on drafts to the assessment of final texts and the academic integrity on which it depends. Below are the three sub-questions that guide the study:

1. Where do EAP instructors locate the opportunities and risks of GenAI within the L2 writing process ?
2. How do EAP instructors perceive GenAI as a source of feedback on L2 writing, and how does their views relate to their own feedback expertise and their overall attitude to GenAI?
3. How do EAP instructors perceive their role of managing academic integrity in AI-assisted L2 writing?

2. Literature Review

2.1 Generative AI across the Stages of L2 Writing

The first wave of research on GenAI in L2 writing was largely conceptual, and much of it was organised around the L2 writing process. Barrot (2024), for instance, proposed an eight-stage model running from preparation and modelling through planning to collaborative writing, with explicit attention to over-reliance, ethics and privacy. The model built on his earlier warning that GenAI tools could weaken students' ownership of their texts, introduce inaccurate content and make assessment harder to trust (Barrot, 2023). In an early classroom study, Su et al. (2023) examined how ChatGPT could augment outlining, revising and reflection in L2 argumentative writing, and they documented problems of accuracy and over-reliance. Such research laid the ground for the idea that GenAI is better understood as a set of stage-specific possibilities within the L2 writing process (Barrot, 2024; Su et al., 2023). What remained unclear was how EAP instructors perceive these affordances and risks of GenAI in their L2 writing classes.

As such, more studies investigated L2 teachers' attitudes towards and perceptions of GenAI integration in L2 writing classes. For example, Barrett and Pack (2023) asked 68 educators and 158 university students to judge the acceptability of GenAI use across six writing tasks. These tasks were brainstorming, outlining, writing, revising, feedback and evaluating. Both educators and students concurred on a lack of preparedness at classroom and institutional levels. However, because respondents rated hypothetical scenarios, the study showed what teachers and students find acceptable in principle rather than what occurs in practice. In the Gulf region, Al Saliti et al. (2025) offered a detailed account in the pre-tertiary education context. Among 340 Saudi secondary-school language teachers, attitudes were neutral at the pre-writing and publishing stages, where AI supports students, but resistant at the drafting and revising stages, where AI produces or rewrites text. Despite the large scale of the study, its secondary-school sample and focus on attitude rather than practice limit its generalisability to university EAP writing programmes.

Research among EAP instructors in higher education shows more nuance. Funabashi (2026), for example, surveyed 45 EAP practitioners at 22 UK institutions and found GenAI in use at every stage of academic writing, most often in planning and revision, with reservations about its use for composing text. Acceptability of integration rose with practitioners' own use of the tools and with how often they discussed AI with students. Similarly, Meniado (2026), drawing on interviews with 14 teachers from nine Southeast Asian countries, found that GenAI was used to produce model texts, support brainstorming and outlining, and provide corrective feedback. These practices, limited in complexity and variety, were attributed to gaps in teachers' technological and pedagogical knowledge (Meniado, 2026). Meniado's (2026) findings are in line with a review of 29 studies of ChatGPT in English as a foreign language (EFL) writing classes, which found that GenAI integration was still at an early stage and centred on revision, proofreading, idea generation and personalised feedback (Le et al., 2025). These studies suggest that L2 writing

teachers use GenAI first where it is easiest to control, while more demanding forms of integration may emerge as their confidence with the tools grows.

In Arabic-speaking contexts, evidence is growing. Al-Azzawi (2026) found that 40 Arab EFL teachers valued AI chatbots for idea generation, linguistic support and initial feedback. The same teachers, however, were concerned about the quality of AI feedback, students' overuse and academic integrity, as well as uneven institutional preparedness and technology infrastructure. Interestingly, they mostly relied on their own mediation, such as staged instruction and verification-based writing activities, to manage these risks. Similarly, Alsuwat (2026) found that, in spite of their generally positive attitude towards ChatGPT's pedagogical value, Saudi university professors stressed their ethical caution and institutional barriers. In Oman, Khattak et al. (2025) found that 21 foundation-programme teachers were more inclined to use AI for preparing instructional materials than for assignments, grading or feedback. Examining GenAI integration more closely, Said Al Fathi's (2026) case study of five teachers in an Omani foundation programme described AI-supported materials and assessment design as an iterative cycle of prompting, checking and adapting. The teachers in that study worried about inaccurate content, cultural inappropriateness, privacy and academic integrity. A subtle distinction can be traced between teacher-facing AI uses, such as preparing materials, and student-facing uses, such as producing text, although it is rarely made explicit. Although growing, the Omani studies have mainly relied on small samples and do not map opportunities and risks across the teaching and learning of L2 writing.

2.2 Generative AI as a Source of Feedback

Feedback is a cornerstone of L2 writing instruction. However, a scoping review of 51 studies found that few examined how teachers use AI in their own feedback practice (Crosthwaite & Sun, 2026). Comparative studies paint a mixed picture of quality. For example, Steiss et al. (2024) found that feedback from trained human evaluators was of higher quality than ChatGPT feedback on most criteria. In the same vein, Pack et al. (2025) found GPT-4 less accurate and less reliable than two experienced educators in giving written corrective feedback on EAP writing. On the other hand, Yu and Xie (2025) found that ChatGPT's surface-level feedback was comprehensive and accurate in a secondary-school setting and stronger than one teacher's, although students took up the teacher's feedback slightly more often. Zou et al. (2025) reported a similar contrast: Chinese undergraduates engaged more with teacher feedback and revised more accurately from it but were particularly successful at using ChatGPT's comments on organisation. Hence, AI feedback has particular strengths and weaknesses that teachers need to recognise when integrating it into their L2 writing classes.

Furthermore, how students engage with AI feedback matters. Zhan and Yan (2025) observed 16 undergraduates working with ChatGPT feedback on IELTS writing tasks. Participants' behavioural engagement was often superficial and their trust in the feedback varied, which suggests that students need guidance in evaluating and acting on AI feedback. Other studies have begun to test hybrid models that combine AI and teacher feedback. Feedback studies have begun to test hybrid, or integrated, models of feedback. For example, Tran (2025) followed 14 Vietnamese undergraduates over 15 weeks and found that combining AI and teacher feedback produced the most revisions. AI feedback tended to address surface-level problems, while teacher feedback addressed higher-order concerns. Interestingly, placing AI feedback first produced slightly more revisions, although the difference between the two orders was not significant. Despite its small sample, the study offers a concrete model of AI-first, teacher-second hybrid feedback.

Two Q-methodological studies have examined L2 teachers' own views on AI feedback (Jinowat et al., 2026; Yao, 2026). Teachers differed considerably in how far they trusted AI feedback, with stances ranging from guided trust to strategic resistance and from pragmatic optimism to ethically cautious critique. Among 36 Turkish university instructors, AI feedback was welcomed as a supplement for linguistic precision. Its interpretive reliability, though, was treated with scepticism. To compensate, teachers filtered AI suggestions and refined their rubrics (Orak, 2025). Similarly, Wei et al. (2026) observed 13 L2 teachers at a research-intensive university over nine weeks. They found that teachers handed language-level feedback to AI more readily but kept control over feedback on content and overall L2 writing quality. Teachers' own experience with AI may also shape these views. For instance, Saeli et al. (2025) revealed that, among 301 Iranian EFL teachers, those who had already used ChatGPT for feedback were more positive about it as a feedback tool. PLD may also shift teachers' beliefs. Rahmati et al. (2025) followed 11 EFL teachers through a PLD programme and found that they developed a more nuanced understanding of how prompts shape ChatGPT feedback. Two of them even reported that reviewing AI feedback sharpened their own feedback literacy. Taken together, teachers' openness to AI feedback appears bound up with experience and PLD as well as with their desire to maintain pedagogical control (Jinowat et al., 2026; Saeli et al., 2025; Wei et al., 2026).

2.3 Academic Integrity, AI Detection and Teacher's Preparedness to Manage Integrity

The weakness of AI detection methods has complicated academic integrity issues. Weber-Wulff et al. (2023) tested 14 detection tools and concluded that none was accurate or reliable. In the same vein, Liang et al. (2023) showed that GPT detectors often misclassify writing by non-native speakers of English as AI-generated. Therefore, researchers have strongly cautioned against using AI detectors to decide integrity cases (Perkins et al., 2024a). Similarly, human judgement is deemed unreliable. Teachers

identified ChatGPT-written essays with 70% accuracy, and their confidence did not predict their accuracy (Waltzer et al., 2023). In experiments at two UK universities, experienced markers could not reliably differentiate between work produced with GenAI and work produced without it (Kofinas et al., 2025).

Survey and interview studies show that L2 writing teachers are concerned about AI-assisted dishonesty but feel ill-equipped to deal with it. Drawing on questionnaires and interviews with 67 EFL teachers, Mohammadkarimi (2023) found that teachers admitted difficulty in detecting AI-generated work and agreed that AI enables dishonesty and hinders the development of writing skills; the teachers called for training and institutional guidelines. In a different study, Ngo et al. (2024) found that Vietnamese EFL teachers attributed AI-related plagiarism to students' lack of ideas, motivation and language competence, and favoured more regulation, detection software and education in responsible use. Among 100 faculty members and graduate students in the Philippines, AI writing tools were valued, but their use was constrained by limited access, a perceived lack of knowledge of how to use them and concerns about integrity (Gustilo et al., 2024). Caribbean EAP lecturers were sceptical of detection software because of false positives and possible bias against multilingual students, preferred assessment redesign, and reported little institutional guidance (Madden et al., 2026). A think-aloud study of eight foundation-level EAP teachers in Bahrain found that they combined their knowledge of L2 students' writing with external tools when judging authorship (Teraif, 2026). Across these contexts, teachers doubt detection tools yet have little institutional guidance on alternatives.

Ethical responsibility and academic integrity of both teachers and learners cannot be separated from institutional policy. In a recent study by Vellanki and Tabassum (2026) in Oman, 52 EFL faculty rated their ethical responsibility and commitment to academic integrity highly ($M = 4.19$) but gave institutional policy the lowest rating of all the dimensions surveyed ($M = 3.25$). They described an "awareness–practice gap" between individual ethics and institutional action. Among students, Omani undergraduates' awareness of institutional policy was negatively associated with their intention to misuse AI (Qoussini et al., 2026), which suggests that policies need to be clearly communicated. A growing body of work argues that institutions should move away from detection towards authentic assessment design. Giray et al. (2025) contended that detection-focused approaches breed distrust among faculty and students and called for process-oriented assessment instead. The AI Assessment Scale (Perkins et al., 2024b), which Roe et al. (2026) adapted for EAP, may help teachers specify what level of AI use is permitted for each L2 writing task.

2.4 Theoretical Lens

The study draws on the technological pedagogical content knowledge (TPACK) framework, which describes the integrated knowledge teachers need in order to teach particular content with technology (Mishra & Koehler, 2006). Subsequently, Mishra (2019) added contextual knowledge to the framework. More recently, Mishra and his colleagues argued that the rise of GenAI systems changes not only the technology component but also what counts as content and pedagogy (Mishra et al., 2023). In the realm of L2 writing instruction, in particular, these GenAI tools are capable of producing the very texts that L2 writing teachers help students learn to write. In response, Celik (2023) extended TPACK to AI with an explicit ethical component, that is, knowing how to use AI-based tools responsibly. Read through this lens, the three research sub-questions concern three facets of teachers' knowledge for AI-mediated L2 writing instruction, which follow the writing process. These are where GenAI fits in the pedagogy of writing from task design to revision (SQ1), how it relates to teachers' expertise in feedback (SQ2), and whether ethical commitment is matched by the knowledge needed to protect the integrity of assessed writing (SQ3).

2.5 Research Gap

It can be inferred from the reviewed studies that L2 teachers judge GenAI differently at different stages of writing, that they are cautiously and conditionally open to AI feedback, and that they face academic integrity challenges that neither AI detection software nor human judgement can resolve. Nevertheless, regional evidence remains scarce. To the best of our knowledge, no study in Oman or the wider Gulf has mapped where university EAP instructors locate the opportunities and risks of GenAI across the stages of L2 writing; the Gulf evidence comes either from secondary schools (Al Saliti et al., 2025) or from small samples focused on materials design (Khattak et al., 2025; Said Al Fathi, 2026). Nor have studies of teachers' views on AI feedback related those views to teachers' own feedback expertise, or examined how EAP instructors' sense of responsibility for academic integrity relates to their preparedness to manage it in AI-assisted L2 writing. This study addresses these gaps with EAP writing instructors in Oman.

3. Methodology

3.1 Research Design

The study adopted an exploratory, cross-sectional survey design combining closed-ended and open-ended items within a single questionnaire. The aim was to map how EAP instructors position generative AI (GenAI) in the L2 writing process. Closed-ended

items were analysed descriptively and correlationally, while open-ended responses were analysed through qualitative content analysis. The data form part of a larger survey of EAP instructors' knowledge and attitudes regarding GenAI integration into L2 writing instruction. This article reports only the items on the L2 writing process, feedback and academic integrity, together with the stage-based coding of the open-ended responses.

3.2 Research Context

The study was conducted among EAP instructors in the general foundation programme and the Department of English of two universities in the Dhofar Governorate, Oman. The foundation programme prepares students who have completed Arabic-medium secondary schooling for English-medium undergraduate study. Academic writing is a core component of its English courses. At the time of data collection, institutional policy required students' written work to be checked with Turnitin. A writing submission was acceptable only if its AI-detection score did not exceed 20%.

3.3 Participants

All 64 English instructors teaching in the foundation programme and the Department of English were invited to respond to the questionnaire. Of these, 51 responded (a response rate of 79.7%). Two respondents were excluded because their response patterns indicated inattention: one selected "Neutral" for all 42 Likert statements, and the other gave an identical response to 37 statements, including reverse-worded ones. Neither provided open-ended answers. The analytic sample was therefore 49, with no missing values on the closed items. Of these 49, 27 were male and 22 female. Most were aged 30–39 ($n = 18$) or 40–49 ($n = 21$), and 25 held a doctoral degree and 22 a master's degree. University teaching experience ranged from under five years ($n = 5$) to more than 20 years ($n = 6$), with most reporting 5–10 years ($n = 15$) or 11–15 years ($n = 15$). Thirty-six taught in the foundation programme, one in an English for Medicine course and 12 in other English courses. Twenty-one (43%) reported having received training on the use of AI in English language teaching. All reported using AI tools in teaching writing at least rarely ($n = 4$), sometimes ($n = 15$), often ($n = 17$) or always ($n = 13$). Table 1 summarises the participants' profile.

Table 1: Profile of participants ($n = 49$)

Characteristic	n	%
Gender		
Male	27	55
Female	22	45
Age (years)		
20–29	3	6
30–39	18	37
40–49	21	43
50–59	4	8
60 or older	3	6
Highest degree		
Doctoral degree	25	51
Master's degree	22	45
Bachelor's degree	2	4
University teaching experience (years)		
Less than 5	5	10
5–10	15	31
11–15	15	31
16–20	8	16

Characteristic	n	%
More than 20	6	12
Teaching context		
Foundation programme	36	73
English for Medicine course	1	2
Other English courses	12	24
Received training on AI in English language teaching		
Yes	21	43
No	28	57
Frequency of AI use in teaching writing		
Rarely	4	8
Sometimes	15	31
Often	17	35
Always	13	27

Note. Percentages are rounded and may not sum to 100.

3.4 Instrument

The questionnaire was administered in English through Google Forms and had three sections:

1. background information (seven items);
2. 42 five-point Likert statements (1 = strongly disagree, 5 = strongly agree), ten of which negatively worded, on teacher knowledge for L2 writing instruction and for integrating AI, organised in nine blocks informed by the TPACK framework and its AI extensions (Celik, 2023; Mishra, 2019; Mishra & Koehler, 2006);
3. an overall attitude rating on a five-point scale (higher scores indicating a more positive attitude) ; and four open-ended prompts.

The items were developed for this study with reference to the TPACK framework and its AI extensions, Intelligent-TPACK and contextual knowledge (Celik, 2023; Mishra, 2019; Mishra & Koehler, 2006). The draft was reviewed by five EAP colleagues, each with ten or more years of experience in teaching academic writing for relevance and clarity. Before the main administration, the questionnaire was piloted with 15 EAP instructors who did not take part in the main study. Pilot participants commented on item clarity, wording and completion time, and the wording of five items was revised. The final version took approximately 10 minutes to complete, a length chosen to limit respondent fatigue and encourage considered answers, particularly to the open-ended prompts.

This article analyses the 17 Likert items listed in Table 2, together with the overall attitude rating, the use-frequency item, the AI-training item and two open-ended prompts: "In your view, what opportunities does the use of AI tools create for English writing teaching in your institute?" (O2) and "What challenges or concerns do you have about integrating AI tools into your English writing teaching?" (O3). The codes follow the order of the open-ended questions in the questionnaire; O1 (comments on overall attitude) and O4 (professional development and policy needs) are not analysed here.

Table 2: *Questionnaire items that were analysed in this study*

Code	Item (as administered)	Focus
L21	I can use AI tools to provide students with appropriate writing samples.	SQ1
L23	I can use AI tools to help my students overcome errors (e.g., word choice) that my students typically make while writing in English.	
L27	I know how to guide my students to use AI tools at each stage of the writing process (planning, drafting, redrafting, and editing).	
L34	I can design writing lessons that effectively integrate AI tools, teaching strategies, and writing content.	

Code	Item (as administered)	Focus
L13	I know how to provide both written and oral feedback on my students' written work.	SQ2
L18	I am able to provide targeted feedback on students' writing content, organisation, and language choice.	
L22	AI enables me to provide better feedback on my students' written work.	
L5	I know how AI detectors work.	
L24	I know how to use tools to detect AI-generated content in my students' written work.	SQ3
L26	I find it difficult to design tasks that prevent students from overusing AI.	
L30	It is challenging for me to design AI-proof English writing tasks.	
L31	I know when my students should or should not use AI in their English writing tasks.	
L35	I know how to communicate clear guidelines to students about the acceptable use of AI tools in English writing tasks.	
L36	I consider the ethical implications of AI-generated content when designing English writing assessments.	
L37	I am unsure how to address academic integrity concerns related to students' AI use in my English writing courses.	
L38	It is my responsibility to check AI-generated content for cultural, religious, and contextual suitability before I use it in teaching English writing.	
L39	It is my responsibility to check if my students use AI to write for them instead of writing by themselves.	

Note. Item codes follow questionnaire order (L1–L42). Agreement with L26, L30 and L37 indicates difficulty or uncertainty.

The three responsibility items (L36, L38, L39) were combined into an ethical-responsibility score (Cronbach's $\alpha = .80$, 95% CI [.68, .88]). The other 14 items were analysed individually because each captures a distinct aspect of practice rather than a shared underlying construct.

3.5 Data Collection and Ethical Considerations

The study received ethical approval from the host university before data collection began. The questionnaire was then distributed electronically via Google Forms and remained open between March and October 2026. Its opening page explained the purpose of the study and stated that participation was voluntary and anonymous. Participants confirmed their consent by agreeing to continue to complete the questionnaire. Participants' identities were anonymised, and respondents are referred to only by number.

3.6 Data Analysis

Quantitative data were analysed using IBM SPSS Statistics (Version 32.0; IBM Corp., 2026). For each item, the percentage of respondents selecting *agree* or *strongly agree* is reported with a Wilson 95% confidence interval, alongside the mean and standard deviation. Associations were estimated with Spearman's rank correlation (ρ), with 95% confidence intervals from 5,000 bootstrap resamples. Paired proportions within respondents were compared with exact McNemar tests: detector use (L24) against understanding of detectors (L5), and ethical responsibility (agreement with both L38 and L39) against integrity confidence (disagreement with L37). Differences between instructors with and without AI training were examined with Mann–Whitney U tests, a non-parametric test suited to ordinal Likert data. Effect sizes for these tests are reported as Cliff's δ (Cliff, 1993), which ranges from -1 to 1 and expresses how often scores in one group exceed those in the other, minus the reverse. Given the exploratory aim and the sample size, p values are reported without correction for multiple comparisons, and interpretation rests on effect sizes and confidence intervals.

Open-ended responses to O2 and O3 were analysed through qualitative content analysis (Schreier, 2012). After content-free replies were removed (for example, "nothing" or "-"), 35 of 39 O2 responses and 39 of 41 O3 responses were coded. A deductive coding frame assigned each response to one or more writing-process stages: *planning*, *drafting*, *revising and editing*, *feedback*, *assessment*, and *teacher preparation* (materials, tasks and model texts). Teacher preparation was included because L2 writing instruction begins with the tasks and model texts that teachers prepare, as reflected in the preparation and modelling stage of Barrot's (2024) model. Responses that did not refer to any particular writing stage (for example, general comments on over-reliance) were coded as "no stage". Segments on feedback (SQ2) and on academic integrity (SQ3) were also extracted. Counts refer to respondents, not mentions. A second coder, an EAP instructor with 15 years of experience in teaching L2 writing who was trained on the coding frame, independently coded all responses. Agreement between the two coders was 90%, and disagreements were resolved through discussion. Quotations are identified by respondent number (R1–R49).

4. Results

4.1 Opportunities and Risks across the Writing Process (SQ1)

EAP instructors located GenAI's opportunities mainly in their own preparatory work and in feedback, and its risks mainly in students' production of text and in assessment (see Figure 1).

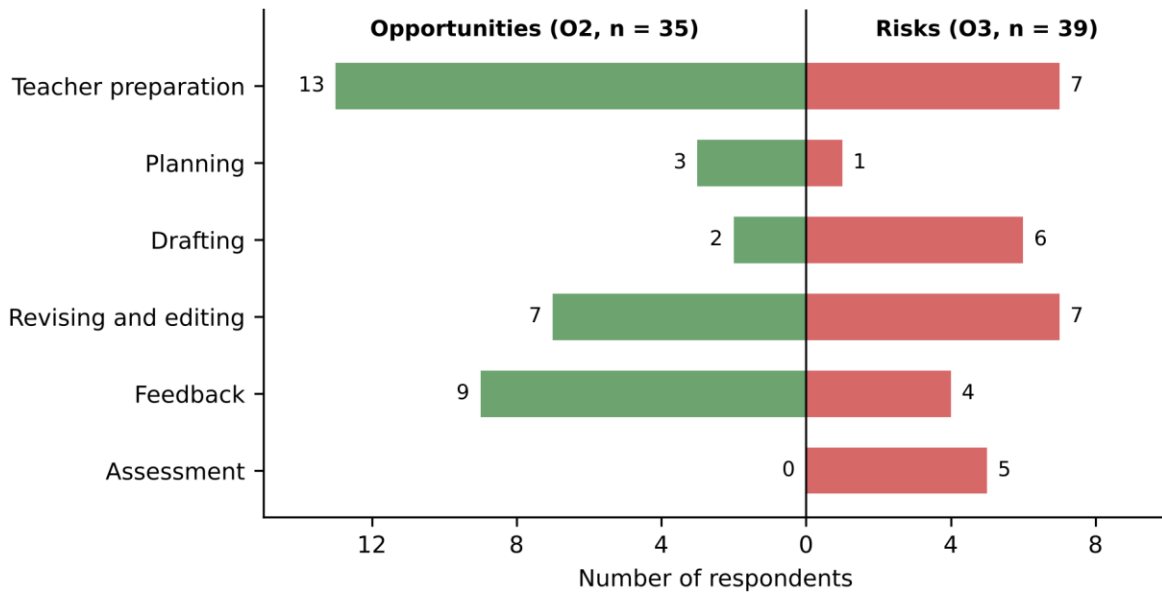


Figure 1: Instructors' location of GenAI opportunities (O2, $n = 35$) and risks (O3, $n = 39$) by L2 writing-process stage, including teachers' preparation of writing materials and model texts

Teacher preparation was the most frequently cited opportunity (13 of 35 respondents). Instructors described generating model texts, writing prompts and tasks differentiated by proficiency level; for example, AI "can provide tailored models appropriate to the students language level" (R12). Feedback was the second most cited opportunity (9 of 35), followed by revising and editing (7), where instructors valued vocabulary, paraphrasing and translation support. One instructor called AI a "good translation tool, paraphrasing helper" that can remove "embarrassment off the students that feel shy" (R38). Planning (3) and drafting (2) were rarely mentioned as opportunities, and AI in assessment was never mentioned as an opportunity.

Risks clustered at different stages. Six respondents described students outsourcing whole texts at the drafting stage ("students tend to outsource their writing to AI", R37). Five raised the authenticity of assessed writing, and seven pointed to revising and editing, where AI "changes a sentence unnecessarily and it confuses the students" (R42). On the teacher-preparation side (7), concerns centred on verifying AI-generated materials: "I always have to review and adapt whatever AI provides" (R12). Interestingly, revising and editing was the only stage cited equally often as an opportunity and a risk (7 and 7). Nine of the 35 opportunity answers and 15 of the 39 challenge answers named no stage, describing GenAI in general terms such as over-reliance or reliability.

Responses to the closed items showed the same teacher-side emphasis (see Table 3). Agreement was higher for using AI to produce writing samples (L21, 71%) than for guiding students' use of AI across writing stages (L27, 59%), helping students address their typical errors (L23, 53%) or designing AI-integrated lessons (L34, 53%).

Table 3: Agreement with writing-process items ($n = 49$)

Item	% agree [95% CI]	M (SD)
L21 Use AI to provide appropriate writing samples	71 [58, 82]	3.96 (1.02)
L27 Guide students' AI use at each writing stage	59 [45, 72]	3.55 (0.96)
L23 Use AI to help students overcome typical errors	53 [39, 66]	3.45 (1.00)
L34 Design writing lessons that integrate AI	53 [39, 66]	3.45 (1.14)

Note. % agree = agree or strongly agree; Wilson confidence intervals.

4.2 GenAI as a Source of Feedback (SQ2)

Instructors were divided on whether AI improves their feedback, and their view was related to their overall attitude towards AI rather than to their confidence as feedback providers (see Table 4).

Table 4: Feedback items and their associations ($n = 49$)

Item	% agree [95% CI]	M (SD)	ρ with L22 [95% CI]
L22 AI enables me to provide better feedback	53 [39, 66]	3.20 (1.17)	–
L13 Know how to provide written and oral feedback	88 [76, 94]	4.14 (0.61)	.10 [–.21, .40]
L18 Able to provide targeted feedback	78 [64, 87]	4.00 (0.79)	.13 [–.19, .42]
Overall attitude towards AI in writing teaching (1–5)	–	3.73 (1.11)	.47 [.21, .67]
The frequency of AI use in teaching L2 writing	–	–	.40 [.14, .63]

Note. ρ = Spearman correlation with bootstrap confidence intervals.

Agreement with L22 (53%) was well below confidence in one's own feedback practice (88% and 78%), and 24% disagreed that AI improves their feedback. Views on AI feedback were unrelated to self-assessed feedback knowledge but moderately related to overall attitude and frequency of use. EAP instructors with and without AI training did not differ significantly on L22 ($M = 3.43$ vs 3.04 ; Mann-Whitney $U = 227$, $p = .152$; Cliff's $\delta = .23$).

In the open-ended answers, the nine instructors who saw feedback as an opportunity stressed immediacy and individualisation, notably in large classes: "Students can get immediate corrective feedback from AI tools on their writings provided that teachers check for plagiarism" (R13). Three participants described a sequence in which AI feedback precedes the teacher's. For example, AI can "provide preliminary feedback. Students can then discuss the output with the instructor" (R41). Four instructors raised feedback-related risks, including students skipping the feedback stage to obtain "flawless responses" (R25), feedback "too general and hard to understand" (R28), dependence on AI reassurance, and a focus on surface corrections.

4.3 Preparedness for Academic Integrity in AI-Assisted Writing (SQ3)

Instructors expressed confidence in setting boundaries for AI use, a strong sense of ethical responsibility, and markedly less confidence in the knowledge and task design needed to act on it (see Table 5).

Table 5: Integrity-related items ($n = 49$)

Item	% agree [95% CI]	M (SD)
L31 Know when students should or should not use AI	84 [71, 91]	4.24 (0.78)
L35 Know how to communicate guidelines on acceptable AI use	82 [69, 90]	4.10 (0.92)
L38 Responsibility to check AI content for cultural and religious suitability	71 [58, 82]	4.00 (0.91)
L24 Know how to use tools to detect AI-generated content	69 [55, 80]	3.94 (1.05)
L39 Responsibility to check whether students use AI to write for them	67 [53, 79]	3.82 (0.91)
L36 Consider ethical implications when designing assessments	65 [51, 77]	3.80 (0.96)
L26 Find it difficult to design tasks that prevent overuse of AI	47 [34, 61]	3.35 (1.18)
L5 Know how AI detectors work	45 [32, 59]	3.39 (1.02)
L30 Find it challenging to design AI-proof writing tasks	39 [26, 53]	3.02 (1.20)
L37 Unsure how to address academic integrity concerns	33 [21, 47]	2.63 (1.44)

Note. Agreement with L26, L30 and L37 indicates difficulty or uncertainty.

Two comparisons of the same instructors' answers to paired items refined this picture. First, more instructors reported knowing how to *use* detection tools than knowing how detectors *work*. Seventeen reported the former without the latter, against five reporting the reverse (exact McNemar $p = .017$). Second, ethical responsibility and integrity confidence did not go together. Of the 29 instructors who endorsed both responsibility items (L38 and L39), 14 (48%) were not confident about addressing integrity concerns. They agreed or were neutral on L37. Overall, 14 instructors were responsible but not confident and 9 confident but not responsible in this sense ($p = .40$). Integrity confidence was unrelated to the ethical-responsibility score ($p = .05$, 95% CI [–.23, .32]) and to detector use (L24; $p = .02$). It was weakly related to finding AI-proof task design less challenging (L30, reverse-scored; $p = .30$, 95% CI [.00, .57], $p = .037$). EAP instructors with and without AI training did not differ in integrity confidence (Mann-Whitney $U = 253$, $p = .400$; Cliff's $\delta = -.14$).

In the open-ended answers, seven instructors raised integrity or assessment validity, ranging from plagiarism ("students might copy and paste what the AI has provided for them", R19) to the evidential value of take-home writing: "if students can generate complete assignments instantly, traditional take-home essays become much less reliable as evidence of learning" (R29).

5. Discussion

5.1 Teacher-Side Value, Student-Side Risk

The clearest pattern in the data is a split between what GenAI does for teachers and what it does to students' L2 writing. EAP instructors saw its value most clearly in their own preparation of materials, model texts and differentiated tasks, and in feedback. They saw its risks most clearly where students produce text and where that text is assessed. This pattern is consistent with Khattak et al.'s (2025) finding that Omani foundation-programme teachers preferred to use AI for preparing materials rather than for grading or feedback. It is also consistent with the iterative materials-design practices that Said Al Fathi (2026) documented in an Omani foundation programme. Furthermore, it echoes UK EAP practitioners' reservations about students using AI to compose text, even though they used it at other stages of writing (Funabashi, 2026). Similarly, teachers in Meniado's (2026) study and in the studies reviewed by Le et al. (2025) used GenAI mainly for model texts, idea generation and proofreading. It appears that teachers begin with AI uses that they can easily check and control.

Similar to the Saudi secondary teachers in Al Saliti et al.'s (2025) study, the EAP instructors in this study saw drafting as a risk, though they differed on revision and feedback. L2 writing revision and editing was contested rather than resisted. On the one hand, instructors valued AI's help with vocabulary, paraphrasing and translation for weaker writers, describing it as a "good translation tool, paraphrasing helper" that can spare shy students embarrassment (R38). On the other hand, they worried that it "changes a sentence unnecessarily and it confuses the students" (R42). Comparative studies of AI and teacher feedback add to this ambivalence. Although AI may address surface-level language well (Yu & Xie, 2025), it is less reliable on higher-order writing concerns (Pack et al., 2025; Steiss et al., 2024). Unlike Al Saliti et al.'s (2025) Saudi teachers, who were neutral about AI at the publishing and feedback-reception stage, the Omani EAP instructors saw feedback mainly as an opportunity. This difference may reflect the university setting, where multi-draft writing assignments make immediate feedback especially useful for revision between drafts.

The closed-response items show that participants were noticeably more confident about using AI to produce writing samples (71% agreement) than about guiding students' use of it across the stages of writing (59%) or designing lessons that integrate it (53%). In TPACK terms, the present data locate the gap in pedagogical rather than technological knowledge: instructors know how to use AI for their own purposes but are less sure how to build students' AI use into writing tasks. Mishra et al. (2023) argue that GenAI makes pedagogical integration newly demanding; the present findings show where, within the L2 writing process, that demand falls, namely at the drafting and revising stages, where student-facing use has to be designed. Using AI for one's own preparation of teaching materials needs little change to established practice, whereas designing how L2 students should use it at each stage of the writing process requires teachers to rethink the writing tasks themselves. Meniado's (2026) finding that GenAI integration practices were limited in variety and complexity suggests that this hesitance is not unique to Oman.

5.2 Feedback Views, Attitude, and Expertise

Just over half of the instructors agreed that AI improves their feedback. However, that view was unrelated to their confidence in giving written and targeted feedback; rather, it tracked their overall attitude to AI and how often they used it. Similar patterns have been reported in previous studies. For example, Saeli et al. (2025) revealed that teachers with prior experience of ChatGPT feedback were more positive about it, and Funabashi (2026) found that EAP practitioners' acceptance of AI integration rose with their own use of the tools. What the present data add is that instructors' confidence in their own feedback (88% and 78% agreement) did not inform how they judged AI feedback. This matters for L2 writing instruction because AI feedback can be accurate on surface features yet miss problems of argument, or offer polished alternatives that do not suit a student's level (Pack et al., 2025; Yu & Xie, 2025). Judging such feedback is precisely where teachers' feedback expertise is needed. If instructors' views of AI feedback rest on general attitude rather than on that expertise, they may adopt or reject it without evaluating it. The AI-first, teacher-second sequence described by three instructors (e.g., R41) resembles the arrangement that produced the most revisions in Tran's (2025) study, but with only three mentions it is better treated as a lead for future research than as a finding.

5.3 Commitment without Capacity

Most EAP instructors accepted responsibility for checking students' AI use (67%) and the suitability of AI-generated content (71%), and most were confident about setting boundaries for AI use (82–84%). This sense of responsibility is unsurprising at an institution whose policy requires Turnitin checks of student writing; as one instructor put it, students may use AI for feedback "provided that teachers check for plagiarism" (R13). However, nearly half of those who accepted responsibility were not confident about handling integrity cases, and integrity confidence was unrelated to the sense of responsibility. In other words, a policy that makes teachers responsible for checking does not by itself prepare them to act. This may reflect Vellanki and Tabassum's (2026) "awareness–practice gap" among Omani EFL faculty, and it supports Celik's (2023) treatment of ethical

knowledge as a distinct component of teacher competence rather than a by-product of ethical commitment. The desire to act responsibly, therefore, should not be conflated with the knowledge needed to do so.

The AI detection finding shows that participants's knowledge concerns how to operate AI detectors rather than how these systems work. In fact, more EAP instructors reported knowing how to use AI detectors than how such detectors work. This gap matters because detection tools are unreliable and biased against non-native writers (Liang et al., 2023; Weber-Wulff et al., 2023). Besides, experienced markers cannot reliably identify AI-assisted work either (Kofinas et al., 2025; Waltzer et al., 2023). EAP instructors who use AI detectors without understanding their limitations and biases therefore risk misjudging the writing of Arabic-L1 students. The risk is sharper at the study site, where institutional policy sets a Turnitin AI-detection threshold of 20%. A numerical cut-off may give detector scores more weight than the evidence on detector accuracy supports. Teraif's (2026) study of foundation-level EAP teachers in Bahrain suggests a better model, in which detection is a matter of professional judgement that combines knowledge of students' writing with cautious use of tools. PLD should therefore focus on helping EAP instructors interpret uncertain detector results.

In the current study, integrity confidence was weakly related to confidence in designing L2 writing tasks that are hard to complete with AI ($p = .30$) but not to AI detector use ($p = .02$). For the teaching of L2 writing, this suggests that instructors' sense of control over integrity comes from how they design writing tasks rather than from the tools they use to check texts. It follows that process-based writing tasks, such as staged drafts, in-class writing and reflections on revision, may serve both learning and integrity, because they make the development of students' texts visible (Giray et al., 2025; Kofinas et al., 2025).

These findings suggest that EAP instructors' views on GenAI integration into L2 writing teaching and learning is uneven. Notably, the 21 instructors who had received AI training did not differ from their untrained colleagues in their views of AI feedback or in integrity confidence, which suggests that the training available to them has not reached the areas where readiness is thinnest. PLD that trains teachers only to operate AI tools is therefore unlikely to suffice. Instead, PLD should build on instructors' existing use of AI for preparing materials and help them design student-facing AI use at each writing stage, judge AI feedback through their own feedback expertise, understand what AI detector scores can and cannot show, and design assessments that make students' writing processes visible. At the institutional level, understanding the limitations of AI detectors calls for revisiting policies that rest on detection thresholds.

5.4 Limitations

The generalisability of these findings is subject to certain limitations. The study draws on self-reported data from EAP instructors at two institutions, so the findings may not generalise to other contexts. Most findings rest on single closed-response items and brief open-ended answers. The associations are correlational and cannot establish direction, and isolated p values near .05 should be cautiously read because the tests were not corrected for multiple comparisons. Participants' responses may also reflect the site's detection-threshold policy, which limits transfer to institutions with different policies. Future research should test these patterns in a multi-institution sample, observe how EAP instructors use GenAI at each L2 writing stage. Further investigations should evaluate PLD aimed at pedagogical design of student-facing AI use, the evaluation of AI feedback, AI detection literacy, and L2 writing assessment redesign.

6. Conclusion

The current study examined how EAP writing instructors in Oman view the role of GenAI within the teaching and learning of L2 writing. EAP instructors placed GenAI's value mainly in their own preparation of instructional materials, model texts and tasks, and in feedback. On the other hand, they placed its risks where students produce text and where that text is assessed. Revising and editing, however, was contested: It was the only stage cited equally often as an opportunity and as a risk. EAP instructors were also divided on whether AI improves their feedback, and their view appears to follow their general attitude to AI rather than their expertise as feedback providers. They were confident in setting boundaries for AI use and felt responsible for integrity. However, that responsibility was not matched by confidence in handling integrity cases. Lastly, more EAP instructors showed knowledge of how to use AI detectors than of how such detectors work.

These findings point to specific targets for PLD and institutional policy in Omani and comparable English-medium programmes. First, because instructors already find value in using AI to prepare materials, PLD can build on this experience and move towards designing students' AI use at each stage of the L2 writing process. Second, PLD should help teachers evaluate AI feedback through their own feedback expertise and teach students how to evaluate and act on AI feedback, thus developing students' feedback literacy. Third, PLD should address the limitations and biases of AI detectors, particularly for L2 writers. Because these tools are unreliable, policies that rely on detection thresholds should be revisited. Finally, since integrity confidence was linked to writing task design rather than to AI detector use, L2 writing assessment should be redesigned so that integrity does not depend solely on AI detection.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: The authors thank the instructors who took part in this study.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- Al Saliti, R. A. M. N., Ismail, A. F., Elmorsy, G. N., & Shahpo, S. M. (2025). Teaching writing skills using generative AI: The paradox of adoption and resistance among language educators. *World Journal of English Language*, 15(8), 28–41. <https://doi.org/10.5430/wjel.v15n8p28>
- Al-Azzawi, I. G. S. (2026). Teachers' perspectives on the challenges of incorporating AI chatbots into writing instruction for Arab EFL learners. *Arab World English Journal*, 17(1), 236–252. <https://doi.org/10.24093/awej/vol17no1.15>
- Al-Raimi, M., Mudhsh, B. A., Al-Yafaei, Y., & Al-Maashani, S. (2024). Utilizing artificial intelligence tools for improving writing skills: Exploring Omani EFL learners' perspectives. *Forum for Linguistic Studies*, 6(2), 1177. <https://doi.org/10.59400/fls.v6i2.1177>
- Alsawat, S. E. (2026). Adoption of ChatGPT in Saudi EFL higher education: Perceptions, institutional readiness, and ethical awareness among university professors. *Arab World English Journal (Special Issue on Artificial Intelligence 3)*, 287–302. <https://doi.org/10.24093/awej/a3.18>
- Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, Article 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Barrot, J. S. (2024). Leveraging ChatGPT in the writing classrooms: Theoretical and practical insights. *Language Teaching Research Quarterly*, 43, 43–53. <https://doi.org/10.32038/ltrq.2024.43.03>
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, Article 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114(3), 494–509. <https://doi.org/10.1037/0033-2909.114.3.494>
- Crosthwaite, P., & Sun, S. (2026). Generative AI and L2 written feedback studies: A scoping review. *RELC Journal*, 57(1), 207–219. <https://doi.org/10.1177/00336882251386530>
- Funabashi, R. L. M. (2026). Exploring the integration of generative AI into EAP academic writing instruction in U.K. higher education. *International Journal of English for Academic Purposes: Research and Practice*, 6(2), 149–183. <https://doi.org/10.3828/ijeap.2026.9>
- Giray, L., Sevnyarayan, K., & Ranjbaran Madiseh, F. (2025). Beyond policing: AI writing detection tools, trust, academic integrity, and their implications for college writing. *Internet Reference Services Quarterly*, 29(1), 83–116. <https://doi.org/10.1080/10875301.2024.2437174>
- Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*, 20, Article 3. <https://doi.org/10.1007/s40979-024-00153-8>
- IBM Corp. (2026). *IBM SPSS Statistics for Windows (Version 32.0)* [Computer software]. <https://www.ibm.com/products/spss-statistics>
- Jinowat, N., Wiboolyasarini, K., Chomchuen, F., & Wiboolyasarini, W. (2026). Trust, resistance, and transformation: A Q-methodological study of teachers' perspectives on AI-generated feedback in second language writing. *Foreign Language Annals*, 59(1), 34–62. <https://doi.org/10.1111/flan.70042>
- Khattak, Z. I., Ghorbanpoor, M., & Mathew, B. P. (2025). Perceptions of EFL teachers on integrating AI tools for foundation-level English language instruction. *International Journal of Education and Practice*, 13(2), 653–668. <https://doi.org/10.18488/61.v13i2.4122>
- Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, 56(6), 2522–2549. <https://doi.org/10.1111/bjet.13585>

- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 5, Article 100156. <https://doi.org/10.1016/j.caeai.2023.100156>
- Le, T. T. P., Quach, T. T. N., Pham, T., & Nguyen, D. L. (2025). A systematic review of how ChatGPT is perceived and utilized in EFL writing classes: Implications for Vietnam. *AsiaCALL Online Journal*, 16(1), 295–311. <https://doi.org/10.54855/acoj.2516115>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), Article 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- Madden, O. N., Smith, R., Edwards, K., & Jackson, J. (2026). Caribbean lecturers' navigation and perceptions of academic integrity and AI-detection software in English for academic purposes (EAP) modules. *Artificial Intelligence in Language Education*, 2, Article 104937. <https://doi.org/10.29140/aile.2026.104937>
- Meniado, J. C. (2026). Teaching L2 writing using generative artificial intelligence: Practices of Southeast Asian language teachers. *Electronic Journal of Foreign Language Teaching*, 23(1), 94–107. <https://doi.org/10.56040/jcmn2312>
- Mishra, P. (2019). Considering contextual knowledge: The TPACK diagram gets an upgrade. *Journal of Digital Learning in Teacher Education*, 35(2), 76–78. <https://doi.org/10.1080/21532974.2019.1588611>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1177/016146810610800610>
- Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and generative AI. *Journal of Digital Learning in Teacher Education*, 39(4), 235–251. <https://doi.org/10.1080/21532974.2023.2247480>
- Mohammadkarimi, E. (2023). Teachers' reflections on academic dishonesty in EFL students' writings in the era of artificial intelligence. *Journal of Applied Learning and Teaching*, 6(2), 105–113. <https://doi.org/10.37074/jalt.2023.6.2.10>
- Mustafa, A., Arbab, A. N., & El Sayed, A. A. (2022). Difficulties in academic writing in English as a second/foreign language from the perspective of undergraduate students in higher education institutions in Oman. *Arab World English Journal*, 13(3), 41–53. <https://doi.org/10.24093/awej/vol13no3.3>
- Naghdipour, B. (2021). English writing pedagogy at the crossroads: The case of Oman. *Journal of Second Language Writing*, 52, Article 100815. <https://doi.org/10.1016/j.jslw.2021.100815>
- Ngo, C., Tran, T., & Nguyen, T. (2024). Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers. *Teaching English with Technology*, 24(1), 28–47. <https://doi.org/10.56297/fsyb3031/mxnb7567>
- Orak, S. D. (2025). Turkish EFL teachers' perspectives on AI-generated feedback: Negotiating trust, control, and pedagogical adaptation in writing instruction. *Applied Linguistics: Research, Measurement and Practice*, 1(1), 74–94. <https://doi.org/10.65334/n40kdq46>
- Pack, A., Hartshorn, K. J., Escalante, J., & Gillette, N. (2025). How well can GenAI (GPT-4) provide written corrective feedback on English-language learners' writing? *International Journal of English for Academic Purposes: Research and Practice*, 5(1), 7–26. <https://doi.org/10.3828/ijeap.2025.2>
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024b). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(06). <https://doi.org/10.53761/q3azde36>
- Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J., & Khuat, H. Q. (2024a). Simple techniques to bypass GenAI text detectors: Implications for inclusive education. *International Journal of Educational Technology in Higher Education*, 21, Article 53. <https://doi.org/10.1186/s41239-024-00487-w>
- Qoussini, A. E., Al Tabib, S., & Freij, A. (2026). Students' perspectives of AI and academic integrity in higher education institutions in Oman. *International Journal of Advanced Computer Science and Applications*, 17(3), 651–655. <https://doi.org/10.14569/ijacsa.2026.0170357>
- Rahmati, P., Hoomanfar, M., & Saeli, H. (2025). Leveraging ChatGPT for L2 writing: Teacher cognition and the impact of professional development. *International Journal of Applied Linguistics*, 36(2), 1653–1664. <https://doi.org/10.1111/ijal.70011>
- Roe, J., Perkins, M., & Tregubova, Y. (2026). The EAP-AIAS: Adapting the AI Assessment Scale for English for academic purposes. *TESOL Journal*, 17(2), Article e70122. <https://doi.org/10.1002/tesj.70122>
- Saeli, H., Rahmati, P., Dalman, M., Shamsi, Y., & Koltovskaia, S. (2025). Embrace or resist? EFL teachers on adopting ChatGPT as a teaching and feedback tool. *Language Awareness*, 35(3), 745–765. <https://doi.org/10.1080/09658416.2025.2577191>
- Said Al Fathi, M. K. (2026). ChatGPT-mediated materials and assessment design in an Omani EFL foundation programme: An Intelligent-TPACK case study. *Language, Technology, and Social Media*, 4(3), 433–447. <https://doi.org/10.70211/ltsm.3026-7196.719>

- Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE. <https://doi.org/10.4135/9781529682571>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, Article 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, Article 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- Teraif, Z. (2026). Who wrote this? An EAP think-aloud study on AI detection. *Language Learning & Technology*, 30(1), 1–22. <https://doi.org/10.64152/10125/73709>
- 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), Article 232. <https://doi.org/10.3390/educsci15020232>
- Vellanki, S. S., & Tabassum, A. T. (2026). EFL teachers' perceptions of ethical AI use in teaching, learning, and assessment: Insights from an Omani university context. *Arab World English Journal*, 17(3), 379–405. <https://doi.org/10.24093/awej/vol17no3.20>
- Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the ability of teachers and students to differentiate between essays generated by ChatGPT and high school students. *Human Behavior and Emerging Technologies*, 2023, Article 1923981. <https://doi.org/10.1155/2023/1923981>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19, Article 26. <https://doi.org/10.1007/s40979-023-00146-z>
- Wei, W., Zhang, X., & Zhu, X. (2026). Embedding, copilot, or agent? L2 teachers' modes of collaboration with generative AI in providing feedback on writing assignments: Rubric influences, workload dynamics, and barriers to integration. *European Journal of Education*, 61(3), Article e70795. <https://doi.org/10.1111/ejed.70795>
- Yao, G. (2026). Exploring teacher perspectives on GPT in L2 disciplinary academic writing through the lens of feedback literacy: A Q-methodology approach. *Assessment & Evaluation in Higher Education*, 51(4), 615–631. <https://doi.org/10.1080/02602938.2025.2553104>
- Yu, H., & Xie, Q. (2025). Generative AI vs. teachers: Feedback quality, feedback uptake, and revision. *Language Teaching Research Quarterly*, 47, 113–137. <https://doi.org/10.32038/ltrq.2025.47.07>
- 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*, 51(5), 821–834. <https://doi.org/10.1080/02602938.2025.2471821>
- Zou, S., Guo, K., Wang, J., & Liu, Y. (2025). Investigating students' uptake of teacher- and ChatGPT-generated feedback in EFL writing: A comparison study. *Computer Assisted Language Learning*, 39(4), 1062–1091. <https://doi.org/10.1080/09588221.2024.2447279>