TESOL Journal

WILEY

CONCEPTUAL FEATURE ARTICLE OPEN ACCESS

The EAP-AIAS: Adapting the AI Assessment Scale for English for Academic Purposes

Jasper Roe[1] | Mike Perkins[2] | Yulia Tregubova[2]

[1]Durham University, Durham, UK | [2]British University Vietnam, Hung Yen, Vietnam

Correspondence: Jasper Roe (jasper.j.roe@durham.ac.uk)

Received: 3 October 2024 | **Revised:** 14 January 2026 | **Accepted:** 16 March 2026

ABSTRACT

The advancement of Generative Artificial Intelligence (GenAI) is complicating traditional techniques in English for Academic Purposes (EAP) instruction. This paper proposes a novel adaptation of an existing scale (the AI Assessment Scale), which is designed to facilitate effective learning and assessment in the era of GenAI. We label this the EAP AI Assessment Scale (EAP-AIAS). The EAP-AIAS consists of five levels, ranging from "No AI" to "Full AI," each delineating appropriate GenAI usage in EAP tasks. This offers a new perspective and approach for educators grappling with how to integrate GenAI into their teaching practice. We discuss the rationale behind this adaptation, considering the unique needs of language learners and the dual focus of EAP on language proficiency and the development of academic skills. By offering a flexible scale, the EAP-AIAS seeks to empower practitioners seeking to deal with the challenge of GenAI integration in EAP and prepare students for an AI-enhanced academic and professional future. This adaptation represents a step towards addressing the pressing need for ethical and pedagogically sound AI integration in language education.

1 | Introduction

Generative artificial intelligence (GenAI) is continuing to complicate teaching and learning in various disciplines. On one end of the spectrum are those who believe that AI and GenAI represent vague, problematic terms for a set of technologies which are fundamentally about commercial gain, rather than improving the educational process, and bring more harm to the practice of English language teaching than good (Caplan 2024). On the other hand, surveys among students and faculty generally reveal positive attitudes toward these technologies, with some caution and concern (Chan and Hu 2023; Roe et al. 2024). Among the general population, surveys demonstrate that many feel that GenAI represents a "world-shifting, permanent, step-change in technology and media," evidenced by the app ChatGPT's ability to garner over 100 million users in as little as 2 months, shattering previous records (Archer 2024, 1).

Despite this attention on both ends of the spectrum, GenAI's role in English for Academic Purposes (EAP) remains a relatively unexplored area (Kostka and Toncelli 2023); however, EAP faces unique and specific challenges in this area due to its dual focus on language proficiency and the development of core academic skills (such as referencing texts, synthesizing knowledge from source material, summarizing, paraphrasing, structuring, and delivering presentations). EAP is at risk as there is the potential for students to use GenAI tools to disguise authorship, complete tasks without genuine engagement, or rely on AI for automated machine translation of texts and even classroom discussion, thereby potentially circumventing the learning processes that EAP aims to facilitate.

Although proctored examinations are also common in EAP, courses often encompass take-home written assignments, portfolios, and presentations as core components of a course, as they are authentic in how they mimic tasks required in later academic studies. These formats, while valuable for developing academic skills, are particularly vulnerable to GenAI-assisted completion or outright AI generation, as tools such as ChatGPT can be used to both fix errors in writing and compose entire

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *TESOL Journal* published by Wiley Periodicals LLC on behalf of TESOL International Association.

essays (Godwin-Jones 2022; Perkins 2023). EAP assessments also evaluate not only the content presented but also the linguistic elements, including grammatical and lexical accuracy, structure, and genre-specific stylistic expressions. Therefore, the ability of GenAI tools to produce grammatically flawless, well-structured, and sophisticated writing poses a significant challenge to this assessment model. Additionally, EAP assessments often target a range of skills that go beyond a student's linguistic ability, such as critical thinking (Calma and Davies 2025; Fan et al. 2024). In a pre-GenAI era, a response essay, for example, would require a student to take several steps involving a certain degree of critical engagement before the actual writing—selecting credible sources, reading and critically engaging with their content, identifying similarities and differences in the views or approaches, synthesizing, paraphrasing, and summarizing, and then creating their own work. With the assistance of GenAI, however, it is possible to proceed from step one directly to the final step, which potentially not only deprives students of valuable critical involvement in the writing process but also intensifies the challenges of measuring skills acquisition and enabling "complex plagiarism" (Perkins et al. 2019); plagiarism that is not easily identified by traditional similarity checker software.

Moreover, in cases where students have been able to use GenAI counter to the guidance of an instructor or institutional guideline, it may be impossible to prove that such usage has taken place. Although GenAI text detection software exists, research indicates that these tools are often inaccurate, easily circumvented (Chaka 2023, 2024; Perkins, Roe, Postma, et al. 2024; Perkins, Roe, Vu, et al. 2024; Weber-Wulff et al. 2023), and potentially biased against speakers of English as an Additional Language (EAL) (Liang et al. 2023). Furthermore, the line between a simple AI-powered writing assistant and an illegitimate overuse that obfuscates authorship is not always clear (Roe et al. 2023). This makes it difficult for educators to reliably identify AI-generated or AI-assisted work. Although there may be cases which are obvious, for example, a low-proficiency English learner submitting a piece of extremely sophisticated, well-written, and grammatically perfect research, with tell-tale phrases characteristic of GenAI writing (Cabanac et al. 2021), this in our estimation is likely to represent a relatively uncommon occurrence.

Another element that draws attention to the need for shared, coherent tools for addressing GenAI use is an evolving discourse in acceptability, which has shifted from attempts to outright ban the use of such tools in education to an acceptance of their value and potential benefits to teaching and learning. In some cases, bans in schools in the US have been hastily repealed as more about GenAI's capabilities and use cases have come to light (Singer 2023). In higher education, as GenAI tools have become increasingly ubiquitous and sophisticated, many universities and academic publishers are adapting their policies to permit their use, provided it aligns with principles of transparency and accountability (Perkins and Roe 2023, 2024). The European Commission and United Nations Educational, Scientific, and Cultural Organization (UNESCO) have also provided guidance documents and policies which emphasize the potential benefits of GenAI in research and education (European Commission 2024; Miao and Holmes 2023). This shift acknowledges the fact that GenAI is unlikely to be a passing fad and instead represents a new and important technology that will be increasingly deployed in society at large.

Much of the current debate on GenAI in language teaching is focused on academic integrity. Concerns have been expressed over the ethical implications of AI-assisted writing, highlighting the potential for novel forms of academic misconduct arising from its inappropriate use (Ngo and Hastie 2025; Şık Keser and Razı 2024) Chan (2025) introduces the term "AI-giarism" to describe the misuse of AI in ways that compromise scholarly originality and transparency. At the same time, academic writing, a core skill within EAP, appears to be undergoing a significant paradigm shift, transitioning from traditional technology-assisted writing practices to more complex models of human–AI collaboration (Nguyen et al. 2024). The above highlights that there is a growing imperative to establish pedagogical frameworks that embed AI literacy into EAP curricula, such as the Artificial Intelligence for Academic Purposes (AIAP) framework proposed by Ngo and Hastie (2025). Such resources must strike a balance between harnessing the potential benefits of these new technologies and teaching the critical evaluation of their uses and limitations. At the same time, it is of vital importance that assessment strategies accurately reflect learners' true abilities and skill development to maintain the integrity and quality of EAP instruction and outcomes.

With this in mind, this paper proposes an adaptation of the AI Assessment Scale (AIAS) (Furze et al. 2024; Perkins, Furze, et al. 2024; Perkins, Roe, and Furze 2024) tailored specifically for EAP contexts. The AIAS is designed to be a flexible, adaptable resource that is suited to multiple disciplines and contexts. However, there are few studies which demonstrate how the AIAS can be applied in practice, and the literature has yet to demonstrate how the AIAS may be used to structure formative or unassessed learning activities in the second language classroom. This is the novel contribution of the current paper, as our adapted scale, the EAP-AIAS engages with the specific types of tasks that are common in EAP specifically, as well as language learning in general. The EAP-AIAS is primarily targeted towards EAP instructors themselves, as it can be used to design formative assessments, in-class activities, or entire schemes of work. However, we view it as a relevant consideration for those involved in curriculum and instructional design as well.

The remainder of this paper is structured as follows: First, we engage with the current literature on EAP, assessment, technology, AI, and GenAI to provide a comprehensive contextual introduction. We then present the adapted EAP AI Assessment Scale (EAP-AIAS), explaining its levels and their applicability to common forms of assessment in EAP. Following this, we discuss the implementation of the EAP-AIAS through a series of exemplar use cases. Finally, we conclude with a discussion of the implications of our findings, acknowledgment of limitations, and proposals for future research directions.

2 | The Evolving Landscape of EAP and the Development of GenAI

Originating in English for Specific Purposes (ESP), EAP has grown to represent a complex set of social practices within both institutional and global contexts (Charles 2012). This evolution reflects

the changing nature of higher education, the internationalization of academia, and the growing recognition of the specific language needs of students engaging in English-medium instruction. Today, the complexity and diversity of EAP contexts encompass a broad spectrum of learners and settings (Hyland and Jiang 2021). Unlike general language acquisition courses, EAP's focus extends beyond solely developing linguistic macroskills (reading, listening, writing, and speaking). Contemporary EAP focuses on the acquisition of the target language for use in specific academic contexts, for example, in lectures, presentations, and seminar discussions, as well as study skills, such as techniques and strategies for note-taking, choosing and evaluating credibility of sources, referencing or using Virtual Learning Environment discussion boards, and may involve English for General Academic Purposes (EGAP) or English for Specific Academic Purposes (ESAP) (Gillett 2022). It may also include teaching elements of culture (Gillett 2022), and navigating culture-specific differences in learning approaches (Roe and Perkins 2020). Additionally, EAP may focus on academic literacies, which involve understanding and producing texts within specific academic discourse communities (Lea and Street 2006), and digital literacies, which are skills for navigating and utilizing digital resources and technologies in academic contexts (Hafner 2019).

AI tools have supported EAP for years, for example, through machine translation used as a scaffold for comprehension and vocabulary during reading-heavy tasks. The recent rise of accessible generative AI systems which produce content directly has intensified attention and practice in EFL and EAP (Pack and Maloney 2023). In this paper we use AI as a broad term that includes both traditional AI based machine translation and grammar checking tools, as well as generative systems. The line between these tools is often blurred in practice, but where the distinction matters, we use GenAI to refer specifically to these generative systems. Areas of practice that require attention when considering the rise of these tools include curriculum design, integrating AI literacy alongside traditional language and academic skills (Ngo and Hastie 2025), assessment methods, developing new approaches to evaluate genuine student work in an AI-assisted environment, and new patterns of human–AI collaboration (Nguyen et al. 2024). Other important areas include pedagogical approaches, adapting teaching methods to leverage AI tools while maintaining a focus on core EAP objectives; and ethical considerations, addressing issues of academic integrity, equity, and the appropriate use of AI in academic contexts.

As EAP is primarily a practical field that focuses on addressing real-life classroom challenges faced by teachers and students (Hyland and Jiang 2021), GenAI represents a challenge that requires actionable, flexible, and practical solutions. There is increasing recognition of the need to enable students to engage in AI-mediated academic communication (Ou et al. 2024) while simultaneously ensuring that the fundamental goals of language acquisition and academic skill development are not compromised.

2.1 | Affordances and Limitations of GenAI in EAP

Understanding the range of potential benefits and risks in integrating AI tools into the EAP context is crucial for developing effective resources for AI use in EAP and identifying the potential affordances for students and educators. GenAI tools, particularly advanced chatbots such as ChatGPT, Gemini, and Claude, can provide learners with opportunities for meaningful practice of the target language. Meaningful practice opportunities potentially enhance learners' overall proficiency (Sha 2009). Research has indicated that interactions with AI chatbots can positively affect language proficiency by drawing attention to form and improving accuracy (Bibauw et al. 2019). Through sophisticated prompting, GenAI chatbots can play various roles in the learning process, acting as teacher, peer, or student, offering a flexible learning experience (Kuhail et al. 2023). Research has also explored the use of prompt engineering with EFL learners to craft human–AI narratives, suggesting a potential innovation in teaching pedagogy (Woo et al. 2024).

GenAI applications can provide rich, varied inputs essential for second language acquisition (SLA). The ability to generate diverse examples and explanations can stimulate interest and contribute to overall language growth (Kohnke et al. 2023), and GenAI tools can create personalized practice materials and facilitate real-time explanations, supporting diverse learning needs and allowing for the creation of tailored exercises that align with individual learner proficiency levels and learning styles. However, the integration of GenAI into EAP presents considerable challenges. There is a risk that students will accept erroneous information produced by GenAI, which is particularly problematic in academic contexts where accuracy and reliability are paramount (Pack and Maloney 2024). Indeed, GenAI tools mislead humans into believing they are intelligent, despite producing inaccurate, made-up, or unreliable content (Zhgenti and Holmes 2023), and at the same time, the information produced by many GenAI tools is by nature culturally biased and representative of a Eurocentric worldview (Roe 2025). This cultural bias could lead to reinforcement of dominant epistemologies in the classroom, while also marginalizing other perspectives from non-Western or nondominant cultural and linguistic traditions. This is especially relevant in the context of EAP, where learners come from a variety of different backgrounds.

AI-generated text or audio content may also be above the proficiency level of many learners, potentially leading to an overreliance on GenAI tools instead of developing their own language skills, which could lead to knowledge loss and a "monoculture" of academic knowledge (Messeri and Crockett 2024, 1). As a result of the quality of outputs from GenAI tools, some have called for the teaching of evaluative judgment regarding GenAI output (Bearman et al. 2024). This also extends to the idea of using GenAI outputs for feedback, and the results of empirical studies suggest that automated writing evaluation from GenAI tools may have a place in English as a New Language (ENL) instruction (Escalante et al. 2023), and that these tools can be used for automated essay scoring with L2 writers (Mizumoto and Eguchi 2023). Dai et al. (2023) found that AI provided feedback was more detailed than that of an instructor, yet showed high levels of agreement with the marking of the instructor. On the other hand, these systems are at an early stage in language assessment (Chiu et al. 2023), and questions remain over automated markings' ethics, legality, cost, and privacy (Kumar 2023). Specifically, little is known regarding the adherence of some GenAI models to data protection requirements,

such as GDPR or the EU AI Act. Educators and instructional designers must carefully consider what they enable or facilitate students to share with models, especially when their personal or sensitive data may be at risk.

2.2 | Implications of GenAI in EAP Assessment

Assessment has become an increasingly important topic in EAP, with bibliometric analysis showing an increasing trend in EAP assessment topics in scholarly publications from 1975 to 2019 (Charles 2022). This growing focus reflects the fact that assessment is a critical component of EAP pedagogy and program effectiveness. EAP assessment is inherently complex because of the intricate nature of language and academic skill evaluation (Bruce 2015). This complexity arises from several factors, including the multidimensional nature of language proficiency, integration of language and content, genre-specific requirements, and cultural and linguistic diversity of EAP students. EAP assessments must evaluate not only linguistic competence but also the ability to use language appropriately in academic contexts, while also requiring students to demonstrate knowledge of academic subjects or research methodologies. Several trends have emerged in EAP assessment in recent years. There is a growing shift from an exam culture to a learning culture, with an increasing emphasis on learner-oriented assessment and more formative and continuous assessment methods (Fazel and Ali 2022). Authentic assessment has gained popularity, with a focus on tasks that reflect real-world academic activities (Uludag and McDonough 2022). Portfolio-based assessments have also seen growing adoption, including a range of tasks such as summarizing texts, creating research proposals, and giving oral presentations (Storch and Tapper 2009). In addition, technology-enhanced assessment has become more prevalent, incorporating digital tools and platforms to allow for more diverse and interactive assessment tasks. As a result, AI literacy will become an increasingly relevant construct, enabling learners to meet immediate course outcomes and develop their preparedness for academic engagement in AI-enabled environments. In light of these contextual factors, we propose that the AIAS can be adapted for EAP to meet these goals and provide a suitable, flexible tool.

3 | The AI Assessment Scale (AIAS): An Adaptation for EAP

3.1 | Overview of the Original AIAS

The AI Assessment Scale (AIAS) was designed by Perkins, Furze, et al. (2024)[1] to help educators in a wide range of disciplines adapt their assessments to the GenAI era. The original scale consists of five levels, ranging from No AI to Full AI, designed to guide students on the appropriate use of GenAI tools while adhering to academic integrity principles. The rationale for these five levels was developed from feedback and development of the original AIAS, its effectiveness being demonstrated by empirical data in reducing academic misconduct cases; results from a pilot study of the AIAS suggest that the use of the scale at an institutional level led to reduced cases of academic misconduct, improved student outcomes, and improved ethical integration of GenAI tools into assessment strategies (Furze et al. 2024). Various adaptations or translations of this scale have also been developed globally (Furze 2024; Kılınç 2024; Roe et al. 2025), demonstrating the potential value of similar tools. In 2024, an updated version of the AIAS was released, which established new levels to account for rapid changes in the technological landscape (Perkins, Roe, and Furze 2024), with a refined peer-reviewed framework subsequently published (Perkins et al. 2025). Anchored in social constructivism and Vygotsky's Zone of Proximal Development (Vygotsky 1978), the AIAS frames GenAI as mediation that can extend learner performance beyond what is reliably achievable independently. In the EAP-AIAS, this is operationalized by sequencing mediation from independent performance (Level 1) to AI-mediated input (Level 2) and guided practice (Level 3), and then to learner-managed use that requires critical evaluation, substantial revision, and strategic integration in complex academic tasks (Levels 4–5). While the original AIAS provides a valuable starting point, the unique needs of EAP necessitate a more tailored approach. EAP's dual focus on language proficiency and academic acculturation (Hyland 2018) requires a framework that prioritizes language and skill acquisition and ensures that the fundamental goal of developing genuine language skills is not compromised by AI use. Furthermore, an adapted AIAS needs to support academic skill development, foster criticality, promote literacy, and provide guidelines on the ethical use of AI tools. To achieve the balance required, we aimed for flexibility over prescriptive criteria, identifying that EAP tasks are highly variable across proficiency levels, institutional settings, and disciplinary contexts. As a result, rather than impose fixed or criterion-driven metrics, we developed the scale through a three-step process, focusing on flexibility and providing a tool that could be adapted and tailored to multiple EAP contexts. Here, "EAP contexts" refers to academic writing, speaking, and multimodal academic tasks. However, the scale is most directly applicable to writing-focused assessment where text generation through GenAI tools can obscure evidence of language production.

This involved:

1. Practitioner collaboration: This involved informal, iterative discussions in our own respective contexts to align the EAP-AIAS with real-world classroom experiences and challenges.
2. Trial-and-error adaptation: We drew on our own existing repertoire of EAP assessment and materials (such as writing tasks and presentations) to evaluate how AI might be used to augment learning activities.
3. Critical reflection and consensus-building: Through discussions and revision of our scale adaptation, we critically assessed its value and explored the pedagogical aims of each AIAS level. Through iterative refinement, we developed a consensus about the items contained in the EAP-AIAS.

3.2 | The EAP AI Assessment Scale (EAP-AIAS)

Our proposed EAP-AIAS builds on the original scale, redefining each level to align it with typical EAP tasks. By explicitly mapping each level to these established EAP approaches, the

EAP-AIAS can provide a more robust and theoretically grounded approach for AI integration in instruction. It is important to note that while there may be significant areas of overlap, each level does have some distinctive features. A Level 4 task may seem similar to a Level 5 task, but a Level 5 task will not require a core focus on critically analyzing and revising outputs; rather, it will be expected that a Level 5 task is selected for an activity when the learners are already capable of engaging in real-time critical judgment and evaluation.

The five levels are nonhierarchical and not sequential. Educators select a level according to task purpose, required evidence of learning, and the amount of GenAI involvement that preserves validity. A course may mix levels across tasks, and adjacent levels can overlap in practice. The levels are presented as context and goal-dependent options where instructors should identify required learning outcomes and constraints first, then choose the level that yields valid evidence. The full EAP AIAS can be seen in Table 1.

We recognize that the broad range of tasks present in EAP assessments means that these identified scale levels might be used at different times and for different purposes throughout EAP education, and that no level of the scale is inherently better or more suitable than another. We have therefore adjusted the original color scheme presented in the AIAS from a Red-Green scale to a more neutral palette of colors. This is to avoid any potential implicit suggestions to educators working with this scale that tasks or assessments with no AI usage are incorrect and that Full AI usage is the most appropriate or target level. At the same time, we highlight that language development is not a linear process; therefore, aligning AI-assisted tasks with EAP instruction will not always follow a progression from L1 to L5. Contextual mapping and analysis of the target environment, the context, and the learner needs must always be the priority. We also understand that program change and development is difficult; the resource constraints of the classroom, the requirements of the institution, and instructor training will all play a role in choosing how, when, and why to use the EAP-AIAS.

To illustrate the practical application of the EAP-AIAS, in alignment with the above, we present a series of exemplar use cases across different EAP contexts below. These tasks may be applied across a range of assessment types, recognizing that formative assessments may offer fantastic opportunities for low-risk experimentation at different levels of the scale.

3.3 | Level 1: No AI Use

In an intensive presessional EAP program for international students, a Level 1 "No AI Use" approach could be applied for the final summative examination. This traditional assessment may be conducted in a secure, invigilated physical environment encompassing listening, reading, speaking, and writing components. By requiring students to demonstrate their EAP skills without the aid of AI or assistive technologies, this approach could provide a clear indication of their ability to independently apply language skills in their future academic studies.

3.4 | Level 2: AI-Assisted Language Input

For first-year undergraduate students in an in-sessional English for Specific Academic Purposes (ESAP) program focusing on business subjects, a Level 2 "AI-Assisted Language Input" approach could be implemented for a formative in-class assessment. GenAI tools may be utilized to produce discipline-specific graded readings, aligning with the EAP-AIAS guidelines for AI-assisted idea generation and structuring. Instructors could guide students to generate texts at appropriate difficulty levels, potentially encouraging them to experiment with the chatbot to explain key terms or interpret meanings outside class time. There are limitations to this approach, as AI-generated readings may contain biases or inaccurate content, and these aspects must be foregrounded with the learners.

3.5 | Level 3: AI for Limited Language Practice

In a postgraduate EAP module focusing on academic writing, a Level 3 "AI for Limited Language Practice" approach could be adopted for a diagnostic assessment of AI supported writing skills. Students might be encouraged to use GenAI applications such as ChatGPT, Claude, or Gemini to develop their sentence and paragraph structure skills in line with the EAP-AIAS guidelines for AI-assisted editing. The process could involve creating discipline-specific word lists, collocations, or clichés based on genre requirements, which could then be incorporated into writing tasks. Students could input their written work into the GenAI application for initial feedback on lexical, grammatical, and structural elements. They might then be tasked with analyzing and justifying any changes made based on AI suggestions, potentially promoting critical engagement with the tool. This Level 3 approach could improve students' use of discipline-specific languages, with the critical analysis of AI grammar suggestions possibly leading to a deeper understanding of linguistic structures.

3.6 | Level 4: AI-Assisted Task Completion With Critical Evaluation

In a postgraduate EAP course on research writing, a Level 4 "AI-Assisted Task Completion with Critical Evaluation" approach could be implemented to develop research proposals representing a summative element of the overall research project. This would align with the EAP-AIAS for AI task completion with human evaluation. The process may begin with an introduction to the UNESCO (Miao and Holmes 2023) and COPE (2023) guidelines on the ethical use of GenAI in research. Students could then use AI tools to generate initial ideas and create outlines for their research proposals. A critical component of this Level 4 task might be the evaluation of AI-generated content, which students would need to revise substantially and justify their changes. Peer review sessions could focus on discussing how effectively students have adapted and improved upon AI-generated content. Final submissions may include a reflection on the process of working with and critiquing AI-generated material, fulfilling the critical evaluation aspect of Level 4.

TABLE 1 | EAP-AIAS.

Level	Description	Focus	Example tasks
Level 1: No AI Use	All language and skills tasks completed without AI assistance.	Developing core language skills and academic competencies independently.	Traditional examinations, in-class presentations, in-class comprehension and critical thinking tasks.
Level 2: AI-Assisted Language Input	AI used to generate or augment input materials.	Enhancing comprehension and analysis skills.	Inviting learners to engage with or create AI-generated texts for reading or listening comprehension and micro skills development.
Level 3: AI for Limited Language or skills Practice	AI used for targeted practice of specific language and discourse features or academic skills development.	Reinforcing particular aspects of language, discourse, academic or discipline-specific conventions.	AI-generated content for controlled or semi-controlled practice of discipline-specific vocabulary and/or discourse features; simulated academic discussions with AI.
Level 4: AI-Assisted Task Completion with Critical Evaluation	Students use AI to assist in complex academic tasks but must critically evaluate and substantially revise AI outputs.	Developing critical thinking and digital literacy alongside language and academic skills.	Using AI for initial research or drafting, followed by substantial human revision, reflection, evaluation and critique.
Level 5: Selective AI Integration for Advanced Skills	More extensive AI use allowed but emphasizing its role in enhancing, not replacing, student work.	Preparing students for real-world academic and discipline-specific scenarios involving AI.	Using AI for data analysis in research projects, developing AI-enhanced academic presentations and discipline-specific outcome tasks

However, such an approach is demanding and time-intensive, making it unsuitable for fast paced courses with a prescribed syllabus.

3.7 | Level 5: Selective AI Integration for Advanced Skills

In an EAP workshop for doctoral students, a Level 5 "Selective AI Integration for Advanced Skills" approach could be employed, representing the highest level of AI integration in the EAP-AIAS. GenAI tools might be used extensively to assist students in synthesizing large volumes of research for a summative literature review. The process might involve prompting AI to summarize key points from multiple papers, followed by a critical evaluation of these summaries. In line with Level 5 of the EAP-AIAS, students could be required to identify gaps in AI-generated syntheses and develop their own arguments based on a more comprehensive understanding of the field. Final presentations may include a critical reflection on the role of AI in the research process and its limitations, potentially demonstrating the advanced skills and critical thinking fostered by the highest level of AI integration. Such an approach is suitable at an advanced level but may lack generalizability to other EAP contexts.

4 | Discussion

The advantage of the AIAS in EAP is that it provides a structured yet flexible approach to GenAI integration, providing clear guidelines for usage in specific tasks. The EAP-AIAS can help empower language educators to effectively incorporate AI tools into their teaching and assessment practices while ensuring that students develop the necessary language and academic skills which can be evaluated through a range of tasks. The EAP-AIAS may be used to help maintain academic integrity by educating students about acceptable use where it is appropriate and providing a system for communication. If students are given a clear scale which clarifies their acceptable use of GenAI and receive encouragement on adhering to values of transparency and ethics, then the EAP-AIAS can help to guide the successful integration of GenAI.

The emphasis on critical evaluation of AI-generated content, particularly in types of tasks relating to Levels 4 and 5, resonates with critical EAP approaches (Benesch 2001), encouraging students to develop evaluative judgment with both content and tools, thus fostering criticality.

It also contributes to the development of metacognitive skills, going beyond higher-order thinking skills and allowing space for developing reflective thinking. In this way, the EAP-AIAS serves as a bridge between traditional EAP instruction and the demands of an increasingly AI-influenced academic environment.

At the same time, the EAP-AIAS must be used sensitively and in a student-centered manner. The student perspective must always be considered first when seeking to implement the EAP-AIAS framework; ideally, such use will include a thorough needs analysis that foregrounds not only what students know about AI and GenAI tools, but what they need to know in order to use them effectively, and their perceptions of AI use in general. Furthermore, consideration needs to be given to potential equity gaps when exploring the use of frameworks such as these. Learners may have differing ranges of digital skill sets, and different levels of awareness, access, and resources to engage with AI, and the same may also go for instructors. Decisions to use certain types of AI technologies in the EAP context must pay attention to these potential drivers of inequity. Where such gaps exist, training programs can help to mitigate them through capacity building for both students and staff, including targeted digital and AI literacy support, shared guidance on ethical use and data privacy, and access arrangements that do not depend on individual ability to pay. In low-resource settings, this can be supported through tool-agnostic task design such as the EAP-AIAS, so assessment expectations remain aligned to learning outcomes rather than tool availability.

Questions remain regarding the long-term impact of AI integration on language and skill acquisition. The principle of maintaining most of the instruction and classroom interaction at Level 1 (No AI Use) appears crucial, but longitudinal studies are needed to confirm its effectiveness. With the influence of AI, academic skills development is likely to go through significant changes. Certain subskills might become obsolete, which will require adaptation of core methodologies and full GenAI integration in the EAP curriculum with the design of a relevant set of assessment tools, possibly AI-powered assessment tools such as EAP Talk—an AI software for assessing academic speaking skills (Wang and Zou 2023).

The issue of academic integrity in AI-enhanced environments remains complex. Our experience suggests that fostering a culture of ethical GenAI use through explicit instruction and reflection is essential. The potential of AI tools to exacerbate existing educational inequalities must be carefully monitored and addressed. The aim is to create a constructive learning environment in an EAP classroom where students benefit from using AI tools in developing their academic skills while learning to use them ethically and maintaining academic integrity. The scale may need further refinement to address discipline-specific needs within EAP, as the appropriate level of AI integration may vary across academic fields and discourses. Given the resource-intensive requirements of assessment, task, and curricula redesign, this may affect the ability to deploy the EAP-AIAS at scale. The current scale may also not fully capture the degrees of AI use in collaborative tasks, an important aspect of academic work that requires further exploration. The fast-paced evolution of GenAI technologies indicates that the scale requires regular updates to remain relevant.

Several avenues for future research emerge from our implementation of the EAP-AIAS. These include conducting longitudinal studies tracking students' language and skills development and academic performance as they progress through the EAP-AIAS levels over extended periods, or by conducting quasi-experimental studies that compare an EAP-AIAS based curriculum to a traditional curriculum. We envisage small-scale, qualitative pilot studies that center on assessment tasks that use varying levels of the AIAS as particularly valuable, along with survey-based studies which evaluate student and

instructor perspectives to continue to iteratively develop the framework.

5 | Conclusion

The EAP-AIAS represents a significant step towards addressing the challenges and opportunities presented by GenAI in EAP instruction. Although further research and refinement are necessary, the scale provides a valuable tool for integrating GenAI tools in a manner that enhances, rather than undermines, the core objectives of EAP. As we move forward, maintaining a balance between embracing technological advancements and preserving the fundamental principles of language acquisition and academic skill development is crucial for shaping the future of EAP pedagogy. This scale offers a structured yet flexible approach to EAP assessment in the GenAI era, empowering both educators and learners to engage critically and ethically with GenAI technologies in academic contexts.

Author Contributions

All authors contributed equally to the generation of the EAP-AIAS and writing up of the manuscript.

Funding

The authors have nothing to report.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

Endnotes

[1] The original AIAS and other adaptations are available to review at www.aiassessmentscale.com.

References

Archer, P. 2024. "What Do People Think of Generative AI?" https://www.bbc.co.uk/mediacentre/2024/what-do-people-think-of-generative-ai/.

Bearman, M., J. Tai, P. Dawson, D. Boud, and R. Ajjawi. 2024. "Developing Evaluative Judgement for a Time of Generative Artificial Intelligence." *Assessment & Evaluation in Higher Education* 49: 1–13. https://doi.org/10.1080/02602938.2024.2335321.

Benesch, S. 2001. *Critical English for Academic Purposes: Theory, Politics, and Practice*. Routledge. https://doi.org/10.4324/9781410601803.

Bibauw, S., T. François, and P. Desmet. 2019. "Discussing With a Computer to Practice a Foreign Language: Research Synthesis and Conceptual Framework of Dialogue-Based CALL." *Computer Assisted Language Learning* 32, no. 8: 827–877. https://doi.org/10.1080/09588221.2018.1535508.

Bruce, I. 2015. "EAP and Assessment." In *Theory and Concepts of English for Academic Purposes*, edited by I. Bruce, 196–209. Palgrave Macmillan. https://doi.org/10.1057/978-1-349-59077-3_12.

Cabanac, G., C. Labbé, and A. Magazinov. 2021. "Tortured Phrases: A Dubious Writing Style Emerging in Science. Evidence of Critical Issues Affecting Established Journals." *arXiv*. arXiv:2107.06751. https://doi.org/10.48550/arXiv.2107.06751.

Calma, A., and M. Davies. 2025. "Assessing Students' Critical Thinking Abilities via a Systematic Evaluation of Essays." *Studies in Higher Education* 51: 1–16. https://doi.org/10.1080/03075079.2025.2470969.

Caplan, A. N. 2024. *The Big AI Lie: "Never Mind the Quality, Feel the Width"*. Nigel Caplan. https://nigelcaplan.com/2024/02/25/the-big-ai-lie-never-mind-the-quality-feel-the-width/.

Chaka, C. 2023. "Detecting AI Content in Responses Generated by ChatGPT, YouChat, and Chatsonic: The Case of Five AI Content Detection Tools." *Journal of Applied Learning and Teaching* 6, no. 2: 2. https://doi.org/10.37074/jalt.2023.6.2.12.

Chaka, C. 2024. "Accuracy Pecking Order—How 30 AI Detectors Stack Up in Detecting Generative Artificial Intelligence Content in University English L1 and English L2 Student Essays." *Journal of Applied Learning and Teaching* 7, no. 1: 1. https://doi.org/10.37074/jalt.2024.7.1.33.

Chan, C. K. Y. 2025. "Students' Perceptions of 'AI-Giarism': Investigating Changes in Understandings of Academic Misconduct." *Education and Information Technologies* 30, no. 6: 8087–8108. https://doi.org/10.1007/s10639-024-13151-7.

Chan, C. K. Y., and W. Hu. 2023. "Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education." *arXiv*. arXiv:2305.00290. https://doi.org/10.48550/arXiv.2305.00290.

Charles, M. 2012. "English for Academic Purposes." In *The Handbook of English for Specific Purposes*, 137–153. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118339855.ch7.

Charles, M. 2022. "EAP Research in BALEAP 1975–2019: Past Issues and Future Directions." *Journal of English for Academic Purposes* 55: 101060. https://doi.org/10.1016/j.jeap.2021.101060.

Chiu, T. K. F., Q. Xia, X. Zhou, C. S. Chai, and M. Cheng. 2023. "Systematic Literature Review on Opportunities, Challenges, and Future Research Recommendations of Artificial Intelligence in Education." *Computers and Education: Artificial Intelligence* 4: 100118. https://doi.org/10.1016/j.caeai.2022.100118.

COPE. 2023. *Authorship and AI Tools*. COPE: Committee on Publication Ethics. https://publicationethics.org/cope-position-statements/ai-author.

Dai, W., J. Lin, F. Jin, et al. 2023. "Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT." *EdArXiv*. https://doi.org/10.35542/osf.io/hcgzj.

Escalante, J., A. Pack, and A. Barrett. 2023. "AI-Generated Feedback on Writing: Insights Into Efficacy and ENL Student Preference." *International Journal of Educational Technology in Higher Education* 20, no. 1: 57. https://doi.org/10.1186/s41239-023-00425-2.

European Commission. 2024. "Guidelines on the Responsible Use of Generative AI in Research Developed by the European Research Area Forum—European Commission." https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/guidelines-responsible-use-generative-ai-research-developed-european-research-area-forum-2024-03-20_en.

Fan, Y., S. Tan, and G. Y. W. Lim. 2024. "EAP Teacher Feedback in the Age of AI: Supporting First-Year Students in EFL Disciplinary Writing | Australian Journal of Applied Linguistics." *Australian Journal of Applied Linguistics* 7, no. 3: 1–31.

Fazel, I., and A. M. Ali. 2022. "EAP Teachers' Knowledge and Use of Learning-Oriented Assessment: A Cross-Contextual Study." *System* 104: 102685. https://doi.org/10.1016/j.system.2021.102685.

Furze, L. 2024. "The AI Assessment Scale in Action: Examples From K-12 and Higher Education Across the World [Blog]." *Leon Furze*. https://leonfurze.com/2024/05/20/the-ai-assessment-scale-in-action-examples-from-k-12-and-higher-education-across-the-world/.

Furze, L., M. Perkins, J. Roe, and J. MacVaugh. 2024. "The AI Assessment Scale (AIAS) in Action: A Pilot Implementation of GenAI Supported Assessment." *arXiv*. arXiv:2403.14692. https://doi.org/10.48550/arXiv.2403.14692.

Gillett. 2022. "What is EAP?" https://www.uefap.com/bgnd/eap.htm.

Godwin-Jones, R. 2022. "Partnering With AI: Intelligent Writing Assistance and Instructed Language Learning." *Language Learning & Technology* 26, no. 2: 2.

Hafner, C. A. 2019. "Digital Literacies for English Language Learners." In *Second Handbook of English Language Teaching*, edited by X. Gao, 899–918. Springer International Publishing. https://doi.org/10.1007/978-3-030-02899-2_46.

Hyland, K. 2018. "Sympathy for the Devil? A Defence of EAP." *Language Teaching* 51, no. 3: 383–399. https://doi.org/10.1017/S0261444818000101.

Hyland, K., and F. Jiang. 2021. "A Bibliometric Study of EAP Research: Who Is Doing What, Where and When?" *Journal of English for Academic Purposes* 49: 100929. https://doi.org/10.1016/j.jeap.2020.100929.

Kılınç, S. 2024. "Comprehensive AI Assessment Framework: Enhancing Educational Evaluation With Ethical AI Integration." *arXiv*. arXiv:2407.16887. https://doi.org/10.48550/arXiv.2407.16887.

Kohnke, L., B. L. Moorhouse, and D. Zou. 2023. "Exploring Generative Artificial Intelligence Preparedness Among University Language Instructors: A Case Study." *Computers and Education: Artificial Intelligence* 5: 100156. https://doi.org/10.1016/j.caeai.2023.100156.

Kostka, I., and R. Toncelli. 2023. "Exploring Applications of ChatGPT to English Language Teaching: Opportunities, Challenges, and Recommendations." *TESL-EJ* 27, no. 3: n3.

Kuhail, M. A., N. Alturki, S. Alramlawi, and K. Alhejori. 2023. "Interacting With Educational Chatbots: A Systematic Review." *Education and Information Technologies* 28, no. 1: 973–1018. https://doi.org/10.1007/s10639-022-11177-3.

Kumar, R. 2023. "Faculty Members' Use of Artificial Intelligence to Grade Student Papers: A Case of Implications." *International Journal for Educational Integrity* 19, no. 1: 1. https://doi.org/10.1007/s40979-023-00130-7.

Lea, M. R., and B. V. Street. 2006. "The 'Academic Literacies' Model: Theory and Applications." *Theory Into Practice* 45, no. 4: 368–377. https://doi.org/10.1207/s15430421tip4504_11.

Liang, W., M. Yuksekgonul, Y. Mao, E. Wu, and J. Zou. 2023. "GPT Detectors Are Biased Against Non-Native English Writers." *arXiv Preprint* arXiv:2304.02819.

Messeri, L., and M. J. Crockett. 2024. "Artificial Intelligence and Illusions of Understanding in Scientific Research." *Nature* 627, no. 8002: 49–58. https://doi.org/10.1038/s41586-024-07146-0.

Miao, F., and W. Holmes. 2023. *Guidance for Generative AI in Education and Research*. UNESCO: United Nations Educational, Scientific and Cultural Organisation. https://unesdoc.unesco.org/ark:/48223/pf0000386693.

Mizumoto, A., and M. Eguchi. 2023. "Exploring the Potential of Using an AI Language Model for Automated Essay Scoring." *Research Methods in Applied Linguistics* 2, no. 2: 100050. https://doi.org/10.1016/j.rmal.2023.100050.

Ngo, T. N., and D. Hastie. 2025. "Artificial Intelligence for Academic Purposes (AIAP): Integrating AI Literacy Into an EAP Module." *English for Specific Purposes* 77: 20–38. https://doi.org/10.1016/j.esp.2024.09.001.

Nguyen, A., Y. Hong, B. Dang, and X. Huang. 2024. "Human-AI Collaboration Patterns in AI-Assisted Academic Writing." *Studies in Higher Education* 49, no. 5: 847–864. https://doi.org/10.1080/03075079.2024.2323593.

Ou, A. W., C. Stöhr, and H. Malmström. 2024. "Academic Communication With AI-Powered Language Tools in Higher Education: From a Post-Humanist Perspective." *System* 121: 103225. https://doi.org/10.1016/j.system.2024.103225.

Pack, A., and J. Maloney. 2023. "Using Generative Artificial Intelligence for Language Education Research: Insights From Using Openai's ChatGPT." *TESOL Quarterly* 57, no. 4: 1571–1582. https://doi.org/10.1002/tesq.3253.

Pack, A., and J. Maloney. 2024. "Using Artificial Intelligence in TESOL: Some Ethical and Pedagogical Considerations." *TESOL Quarterly* 58, no. 2: 1007–1018. https://doi.org/10.1002/tesq.3320.

Perkins, M. 2023. "Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era: ChatGPT and Beyond." *Journal of University Teaching & Learning Practice* 20, no. 2: 1–24. https://doi.org/10.53761/1.20.02.07.

Perkins, M., U. Basar Gezgin, and R. Gordon. 2019. "Plagiarism in Higher Education: Classification, Causes and Controls." *Pan-Pacific Management Science* 2: 3–21. https://doi.org/10.13140/RG.2.2.20694.11841.

Perkins, M., L. Furze, J. Roe, and J. MacVaugh. 2024. "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, no. 6: 6. https://doi.org/10.53761/q3azde36.

Perkins, M., and J. Roe. 2023. "Decoding Academic Integrity Policies: A Corpus Linguistics Investigation of AI and Other Technological Threats." *Higher Education Policy* 37: 633–653. https://doi.org/10.1057/s41307-023-00323-2.

Perkins, M., and J. Roe. 2024. "Academic Publisher Guidelines on AI Usage: A ChatGPT Supported Thematic Analysis [Version 2; Peer Review: 3 Approved, 1 Approved With Reservations]." *F1000Research* 12: 1398. https://doi.org/10.12688/f1000research.142411.2.

Perkins, M., J. Roe, and L. Furze. 2024. "The AI Assessment Scale Revisited: A Framework for Educational Assessment." *arXiv Preprint* arXiv:2412.09029.

Perkins, M., J. Roe, and L. Furze. 2025. "Reimagining the Artificial Intelligence Assessment Scale (AIAS): A Refined Framework for Educational Assessment." *Journal of University Teaching and Learning Practice* 22, no. 7. https://doi.org/10.53761/rrm4y757.

Perkins, M., J. Roe, D. Postma, J. McGaughran, and D. Hickerson. 2024. "Detection of GPT-4 Generated Text in Higher Education: Combining Academic Judgement and Software to Identify Generative AI Tool Misuse." *Journal of Academic Ethics* 22: 89–113. https://doi.org/10.1007/s10805-023-09492-6.

Perkins, M., J. Roe, B. H. Vu, et al. 2024. "Simple Techniques to Bypass GenAI Text Detectors: Implications for Inclusive Education." *International Journal of Educational Technology in Higher Education* 21, no. 1: 53. https://doi.org/10.1186/s41239-024-00487-w.

Roe, J. 2025. "Generative AI as Cultural Artifact: Applying Anthropological Methods to AI Literacy." *Postdigital Science and Education* 7: 1107–1124. https://doi.org/10.1007/s42438-025-00547-y.

Roe, J., and M. Perkins. 2020. "Learner Autonomy in the Vietnamese EAP Context." *Asian Journal of University Education* 16, no. 1: 13–21.

Roe, J., M. Perkins, and L. Furze. 2025. "From Assessment to Practice: Implementing the AIAS Framework in EFL Teaching and

Learning." *arXiv*. arXiv:2501.00964. https://doi.org/10.48550/arXiv.2501.00964.

Roe, J., M. Perkins, and D. Ruelle. 2024. "Is GenAI the Future of Feedback? Understanding Student and Staff Perspectives on AI in Assessment." Intelligent Technologies in Education. https://open-publishing.org/journals/index.php/ited/article/view/1184.

Roe, J., W. A. Renandya, and G. M. Jacobs. 2023. "A Review of AI-Powered Writing Tools and Their Implications for Academic Integrity in the Language Classroom." *Journal of English and Applied Linguistics* 2, no. 1: 3.

Sha, G. 2009. "AI-Based Chatterbots and Spoken English Teaching: A Critical Analysis." *Computer Assisted Language Learning* 22, no. 3: 269–281.

Singer, N. 2023. "Despite Cheating Fears, Schools Repeal ChatGPT Bans." *The New York Times*. https://www.nytimes.com/2023/08/24/business/schools-chatgpt-chatbot-bans.html.

Şık Keser, K., and S. Razı. 2024. "Upholding Academic Integrity in EAP." In *Teaching English for Academic Purposes: Theory Into Practice*, edited by E. Zehir Topkaya and H. Çelik, 321–342. Springer Nature. https://doi.org/10.1007/978-3-031-72545-6_14.

Storch, N., and J. Tapper. 2009. "The Impact of an EAP Course on Postgraduate Writing." *Journal of English for Academic Purposes* 8, no. 3: 207–223. https://doi.org/10.1016/j.jeap.2009.03.001.

Uludag, P., and K. McDonough. 2022. "Exploring EAP Instructors' Evaluation of Classroom-Based Integrated Essays." *TESOL Journal* 13, no. 3: e670. https://doi.org/10.1002/tesj.67.

Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Vol. 86. Harvard University Press.

Wang, C., and B. Zou. 2023. "Review of EAP Talk." *Journal of China Computer-Assisted Language Learning* 4: 376–385. https://doi.org/10.1515/jccall-2023-0019.

Weber-Wulff, D., A. Anohina-Naumeca, S. Bjelobaba, et al. 2023. "Testing of Detection Tools for AI-Generated Text." *International Journal for Educational Integrity* 19, no. 1: 1. https://doi.org/10.1007/s40979-023-00146-z.

Woo, D. J., K. Guo, and H. Susanto. 2024. "Exploring EFL Students' Prompt Engineering in Human–AI Story Writing: An Activity Theory Perspective." *Interactive Learning Environments* 33: 1–20. https://doi.org/10.1080/10494820.2024.2361381.

Zhgenti, S., and W. Holmes. 2023. "Generative AI and Education: Adopting a Critical Approach." https://discovery.ucl.ac.uk/id/eprint/10177781/1/Zhgenti%20and%20Holmes%20-%202023%20-%20Generative%20AI%20and%20Education%20Adopting%20a%20Critical%20A.pdf.