Pegem Journal of Education and Instruction, Vol. 16, No. 1, 2026 (pp. 572-587)

RESEARCH ARTICLE WWW.PEGEGOG.NET

Co-Constructing AI Policies for Self-Assessment in English for Specific Purposes: A Feasibility and Needs Analysis Study

Leila Aouam [1] , Leila Benstaali [2]

[1] English Language Department , Abdelhamid Ibn Badis University ,Mostaganem ,Algeria. Email: Leila.aouam@etu.univ-mosta.dz

[2] English Language Department , Abdelhamid Ibn Badis University ,Mostaganem ,Algeria . Email: leila.benstaali@univ-mosta.dz

Received: 13/05/2025 ; Accepted: 28/11/ 2025 ; Published: 06/01/2026

Abstract

Artificial intelligence tools offer potential solutions to persistent feedback challenges in English for Specific Purposes (ESP) education, yet institutional responses have oscillated between ineffective prohibitions and unguided adoption. This feasibility and needs analysis study investigates whether co-constructed AI usage policies(developed collaboratively by ESP learners and trainers) constitute a viable alternative to restrictive bans. Drawing on mixed-methods data from 22 ESP students and in-depth interviews with three experienced trainers across Business English, Occupational English, and Academic English contexts, we examine stakeholder attitudes toward AI-enabled self-assessment, concerns about academic integrity and skill development, and receptivity to participatory policy development. Findings reveal significant contextual variation in how trainers experience delayed feedback and perceive AI viability, ranging from pragmatic openness (Business English) to capacity-constrained skepticism (Academic English) to equity-based resistance (Occupational English). ESP students demonstrate sophisticated understanding of appropriate AI boundaries, strong preference for co-constructed rules over bans or instructor-only policies, and confidence in their capability for participatory engagement. We conclude that co-construction is educationally sound but institutionally demanding, requiring explicit facilitation support, workload adjustments, and foundational infrastructure investment to achieve genuine participatory processes. The study contributes to educational technology governance scholarship by demonstrating that effective AI integration emerges through sustained dialogue among affected stakeholders rather than unilateral mandate.

Keywords: artificial intelligence, English for Specific Purposes, participatory policy development, feedback, self-assessment.

1. Introduction

Artificial intelligence (AI) is rapidly reshaping language education, offering new forms of support while also raising pedagogical and governance challenges (Huang et al., 2024). In English for Specific Purposes (ESP), where learners must develop discipline-specific communicative competence, timely and contextualized feedback is crucial for progress and motivation (Plastina, 2015; Taylor, 2024). Generative AI tools have therefore attracted attention as potential feedback and self-assessment aids because they can provide immediate responses at scale, potentially supporting iterative revision and self-regulated learning (Alshehri, 2025).

Yet institutional responses to generative AI often fluctuate between prohibition and informal, unregulated use, leaving educators and students unclear about appropriate boundaries (Huang et al., 2024). This ambiguity is especially consequential in ESP, where feedback delays and generic guidance can be particularly limiting given the need to develop specialized vocabulary, genre awareness, and professional discourse conventions.

Despite increasing discussion of AI in education, three gaps remain insufficiently addressed in ESP-focused research. First, stakeholder-first governance approaches are widely advocated, but empirical work on genuinely participatory policy development in ESP classrooms remains limited (Partnership on AI, 2023; Fuligni et al., 2025). Second, delayed and generic feedback has not been examined thoroughly as a primary pedagogical justification for AI-supported self-assessment in ESP, including how constraints may vary across proficiency levels and task types (Plastina, 2015; Taylor, 2024). Third, there is still limited evidence comparing how learners and trainers evaluate co-constructed AI policies versus bans or instructor-only rules in authentic ESP settings.

Accordingly, this exploratory research investigates the feasibility, desirability, and design parameters of co-constructed AI policies for self-assessment in ESP learning environments. It examines trainers' experiences of the feedback-delay problem and their views of AI-enabled self-assessment (RQ1), learners' attitudes toward AI use and related integrity/skill-development concerns (RQ2), and both groups' perspectives on co-construction compared with prohibition or instructor-only approaches (RQ3).

2. Literature Review

To understand the feasibility and desirability of co-constructed AI policies in ESP education, this review examines five interconnected bodies of scholarship: (1) evidence on why blanket prohibitions fail as governance mechanisms; (2) contemporary scholarship advocating participatory approaches to AI policy development; (3) research documenting persistent feedback challenges in ESP contexts and how these challenges differentially affect learners; (4) empirical evidence on AI's capabilities and limitations as a feedback solution; and (5) frameworks for designing participatory policy development that centers stakeholder voices. Synthesizing these literatures reveals that while AI-mediated feedback offers genuine solutions to temporal and specificity constraints in ESP teaching, effective implementation requires governance structures honoring learner agency and contextual adaptation rather than universal mandates. This theoretical foundation shapes the research questions and methodology adopted to conduct the study .

2.1 The Case Against Restrictive AI Policies in Education

Evidence from across higher education demonstrates the impracticality and counterproductivity of blanket bans on generative AI tools. Huang et al. (2024) analyzed ChatGPT policies at 500 top universities, revealing that only 32.5% had implemented formal policies by mid-2023. More telling than this policy vacuum is how institutions that did establish policies predominantly chose adoption over restriction: 67.4%

embraced ChatGPT while 32.6% enacted bans. This disparity becomes even more significant when considering that early adopters of restrictive policies frequently reversed their bans after recognizing that students readily accessed these tools through non-institutional channels, effectively rendering prohibitions obsolete.

Three evidence-based considerations justify moving beyond blanket bans on AI. First, complete prohibition denies students and educators opportunities to develop AI literacy competencies essential for future professional contexts (Huang et al., 2024; Moorhouse & Kohnke, 2024). Second, such policies create inequitable conditions, advantaging students with unrestricted home access while disadvantaging those reliant on institutional resources (Partnership on AI, 2023). Third, restrictive approaches prevent systematic exploration of AI's documented pedagogical benefits including personalized feedback, adaptive learning pathways, and support for learner heterogeneity (Alshehri, 2025; Huang et al., 2023). Rather than enforcing prohibition, institutions increasingly adopt governance structures that guide responsible AI usage while preserving opportunities for learning through technology (Huang et al., 2024).

2.2 Toward Participatory Policy Frameworks

Contemporary scholarship increasingly advocates stakeholder-first approaches to AI governance, positioning affected individuals as essential partners rather than subjects of external mandates (Partnership on AI, 2023). The Guidelines for Participatory and Inclusive AI emphasize that "meaningful collaboration between AI practitioners and stakeholders from socially marginalized identities and communities" must inform system development,a principle with particular relevance for education, where it positions learners and trainers as partners in policy formulation (Partnership on AI, 2023, p. 2). This participatory vision receives theoretical grounding from frameworks like the Contextualized Perceptions for the Adoption of LLMs in Education (Co-PALE), developed by Fuligni et al. (2025), which operationalizes stakeholder-first methodology by prioritizing the goals, contexts, and perceptions of educational agents. However, research on participatory AI development reveals a persistent gap: 67% of documented AI projects consulted stakeholders for input but stopped short of transferring genuine decision-making authority (Fuligni et al., 2025), suggesting that authentic co-construction,where learners exercise substantive agency over policy parameters,remains underexplored. Wilson et al. (2025) addressed this gap through the Co-Design Stories Toolkit, demonstrating how participatory design methodologies enable stakeholders to articulate their values and preferences for AI systems affecting their education. Through structured dialogue, such toolkits surface contextual requirements that technical designers might otherwise overlook, enabling more responsive system design.

2.3 Feedback Challenges in ESP: Delayed Response and Generic Content

ESP instruction confronts two interconnected challenges in feedback provision: temporal delays and insufficient disciplinary specificity. Traditional teacher-centered feedback practices suffer from delays that significantly diminish learning impact (Plastina, 2015; Taylor, 2024). Research documents that external corrective feedback exhibits "inhibiting and discouraging effects on learning" with negative impacts on learner affect (Plastina, 2015, p. 39), partially attributable to delays between student performance and feedback reception. When students submit specialized discourse tasks, instructor workload necessitates extended waiting periods before evaluative responses, a constraint particularly severe in large classroom contexts.

Beyond temporal limitations, feedback in ESP contexts frequently lacks disciplinary specificity essential for meaningful language development. ESP pedagogy demands domain-specific linguistic precision that general English feedback cannot adequately address (Plastina, 2015). Students pursuing English for Economics, Medical English, or Business English require feedback attuned to specialized vocabulary, genre conventions, and professional communicative norms specific to their target disciplines. Research consistently documents that teacher feedback in ESP contexts often remains insufficiently responsive to learners' disciplinary and proficiency needs (Stanišić et al., 2022). Evidence from Medical English contexts illustrates this gap: traditional external feedback frequently fails to address the nuanced requirements of medical discourse, prompting exploration of alternative assessment mechanisms including self-generated feedback (Plastina, 2015).

These challenges manifest differentially across proficiency levels, creating inequitable learning experiences within mixed-proficiency cohorts. Higher-proficiency students actively request more substantive feedback and express frustration that course content remains "too easy" without advancing them toward more sophisticated competencies (Taylor, 2024, p. 131). Lower-proficiency learners, by contrast, struggle with motivation during extended feedback delays. In large ESP classrooms with heterogeneous proficiency distributions, feedback provision inevitably skews toward addressing common errors at median levels, systematically disadvantaging both advanced students requiring sophisticated guidance and struggling students needing more frequent, immediate corrective response (Stanišić et al., 2022).

2.4 AI-Enabled Immediate Feedback as Partial Solution

Artificial intelligence technologies address temporal and specificity limitations inherent in human-only feedback systems (Alshehri, 2025; Huang et al., 2023). AI-powered platforms provide immediate responses to learner performance, enabling formative assessment loops that support iterative improvement. Research documents that AI feedback provision associates positively with perceived language proficiency improvement, with learners rating AI feedback as "accurate and, in many cases, more useful than human-generated feedback due to its comprehensiveness and coverage" (Alshehri, 2025, p. 8).

In ESP contexts, AI tools offer domain-specific feedback calibrated to specialized discourse requirements. AI systems generate "targeted linguistic input aligned with learners' professional communication needs," facilitating "access to international academic networks and authentic materials" while "enabling effective feedback and assessment related to both content and language skills" (Assassi, 2025, p. 5). The capacity for AI to analyze specialized vocabulary, genre conventions, and register appropriateness provides ESP learners with feedback granularity that generalist instructors may struggle to deliver consistently.

However, AI feedback systems exhibit important limitations requiring human instructor involvement. Research consistently documents concerns about AI reliability, data privacy, potential over-reliance, and limited enhancement of creativity and critical thinking skills. Higher-proficiency ESP students found AI feedback insufficient for advanced language development needs, acknowledging that "the instructor is still useful...AI tools could help check grammar, but the instructor's advice is still better for content discussion and academic writing style" (Taylor, 2024, p. 132). Similarly, LLM-generated feedback did not improve self-assessment accuracy on average, with effectiveness depending on students' initial accuracy rather than performance levels (Liebenow et al., 2025).

The concept of "human-AI orchestration" emerges as essential for effective ESP pedagogy. Effective integration positions AI as providing immediate, personalized formative feedback while instructors offer contextualized guidance, complex discourse evaluation, and socio-emotional support that automated systems cannot replicate (Taylor, 2024). This complementary approach balances AI's affordances with human expertise essential for deep learning outcomes.

2.5 Participatory Design and Co-Construction Frameworks

Participatory design approaches offer structured methodologies for involving ESP learners and trainers in AI policy development. Research demonstrates that mixed-methods processes combining surveys with follow-up interviews surface context-specific requirements that purely technical approaches would miss (Nadarzynski et al., 2025). This finding suggests that authentic stakeholder involvement yields more responsive and contextually appropriate outcomes.

Educational technology benefits from design thinking frameworks adapted for policy co-creation. Educators have developed protocols for co-creating AI policies with students, emphasizing empathy, ideation, prototyping, and testing (Wilson et al., 2025). The process begins with empathy interviews where students share learning experiences and perspectives on AI's role in their academic lives, identifying themes and concerns grounded in lived experience. From these empathy insights, stakeholders collaboratively formulate guiding questions that frame policy objectives, ensuring solutions resonate with learners' lived experiences and aspirations rather than imposing externally defined parameters.

The EAP-AIAS framework developed by Roe (2024) demonstrates how tiered approaches guide AI integration decisions across diverse assessment contexts. The scale ranges from "No AI" to "AI Exploration," with each level delineating appropriate AI usage in EAP tasks. This flexible framework acknowledges that ESP courses encompass diverse assessment types—writing tasks, oral presentations, research projects—each warranting different AI integration parameters. By providing guidelines rather than rigid prescriptions, the framework enables contextual adaptation responsive to specific ESP disciplinary needs, learner proficiency levels, and institutional contexts.

Effective stakeholder engagement adheres to principles ensuring authentic participation. Partnership on AI's guidelines establish that "stakeholders should determine the manner and extent of their involvement at each design stage," with "participation driven by participants' interest and ability rather than researcher aims" (Partnership on AI, 2023, p. 5). This principle accommodates diverse participation forms rather than imposing uniform engagement requirements. Transparency constitutes another essential principle: clear communication about how learner and trainer input influences policy parameters and acknowledgment of technical or institutional constraints contribute to perceived legitimacy (Chaudhry et al., 2024). Co-learning represents a third critical principle, facilitating "two-way sharing of information and background to allow meaningful contribution and conversation," using "written or visual materials" and "avoiding jargon" to ensure accessibility (Partnership on AI, 2023). Together, these principles establish conditions for authentic stakeholder voice in policy development.

3. Methodology

3.1 Research Design

This study employs a convergent parallel mixed-methods design, combining quantitative questionnaire data from ESP learners with qualitative interview data from ESP trainers. This approach enables concurrent collection and analysis of both data types, with subsequent integration providing comprehensive insights into research questions. The convergent design permits validation of findings across methodological approaches while allowing qualitative depth to contextualize quantitative patterns.

3.2 Participants and Setting

The study recruited 22 ESP students (*M* age = 22.3 years, *SD* = 3.1) from undergraduate and postgraduate programs at an Algerian university and a professional training center respectively, encompassing Business English, Medical English, Engineering English, and Academic English specializations. Participants reported proficiency levels between B1–C2 (CEFR), with exposure to regular instructor feedback on written or oral work. Stratified convenience sampling ensured disciplinary representation across ESP specializations.Besides, three experienced ESP instructors participated in semi-structured interviews. Participants included: (1) a Business English trainer with 8 years experience, managing 90 students across three sections; (2) an Occupational English trainer with 6 years experience, teaching 32 working professionals; (3) an Academic English trainer with 12 years experience, supporting 120 undergraduate and postgraduate students. Purposive sampling recruited maximum variation across teaching experience, specializations, and perspectives on AI integration.

3.3 Data Collection Instruments

For data collection ,two tools are adopted that suit the mixed method approach. Learner Questionnaire : a 61-item structured questionnaire across five sections measured student experiences and attitudes using five-point Likert scales (1 = strongly disagree; 5 = strongly agree). Section A assessed demographics (age, gender, academic level, proficiency, prior AI experience). Section B examined delayed feedback experiences and impacts. Section C measured attitudes toward AI for self-assessment. Section D assessed receptivity to co-construction. Section E included open-ended reflections. The questionnaire required approximately 15-20 minutes to complete". Reliability testing (Cronbach's α = .78–.86 across scales) indicated adequate internal consistency.

While trainer interview is a semi-structured interview (25–35 minutes) explored: (1) current feedback processes and workload constraints; (2) perceptions of delayed feedback impacts; (3) attitudes toward AI-enabled self-assessment; (4) reactions to co-construction approaches; (5) feasibility concerns and implementation requirements; (6) comparisons of co-construction versus bans. Interviews were conducted in-person .

3.4 Data Analysis

Likert scale responses underwent analysis (Quantitative Analysis)using descriptive statistics (frequencies, percentages, means, standard deviations) and inferential statistics. While qualitative responses in the questionnaire received inductive thematic analysis identifying recurring themes.

Similarly , interview transcripts underwent reflexive thematic analysis in iterative phases: (1) familiarization through repeated reading; (2) initial coding assigning descriptive codes to meaningful segments; (3) candidate theme development grouping related codes; (4) theme refinement evaluating coherence and distinction; (5) definition and naming of final themes with illustrative quotations; (6) narrative reporting integrating themes with verbatim quotes.

Integration. Quantitative and qualitative findings were integrated through convergence analysis (identifying alignment), divergence analysis (exploring tensions), complementarity analysis (using qualitative findings to contextualize quantitative patterns), and joint display techniques presenting findings side-by-side for direct comparison.

3.5 Ethical Considerations and Validity

The study received ethics committee approval prior to data collection. Learners provided informed consent by completing questionnaires (with oral explanations); trainers approved orally to the ethical guides explained to them with assurances of voluntary participation and right to withdrawal. All data employed anonymous codes or pseudonyms, with identifying details redacted. Electronic data were password-protected; physical documents secured in locked storage.

Quantitative validity was ensured through instrument grounding in prior ESP literature, pilot testing for clarity, and internal consistency checks. Stratified sampling enhanced representativeness across ESP specializations. Qualitative trustworthiness was established through credibility (prolonged engagement, triangulation with questionnaire data, peer debriefing), transferability (thick description of context), dependability (detailed audit trails), and confirmability (reflexive journals, negative case analysis, transparent quotation use).

4. Findings

4.1 Research Question 1: Trainer Perceptions of Delayed Feedback and AI Viability

4.1.1 Contextual Variation in Feedback Delay Experience

Trainers experienced delayed feedback as contextually variable problems. The Business English trainer (BET) reported 10–14 day turnaround with 90 students across three sections, describing delayed feedback as "a concrete, solvable problem affecting learning effectiveness." He noted specific repeated errors (e.g., confusion between "recommend" and "recommend to") that students failed to correct in subsequent assignments due to temporal distance from feedback. The delayed feedback problem prevented expansion of writing practice: "I can't assign more writing because I can't manage feedback in reasonable time."

In contrast, the Academic English trainer (AET) managed 120 students with 15–20 hours/week grading, experiencing delayed feedback as "symptomatic of systemic resource constraints beyond individual control." She emphasized that feedback delays prevented "expanded writing assignments," forcing her to limit writing practice below pedagogically ideal levels. Yet unlike the Business trainer who experienced this as a manageable problem, the Academic trainer viewed it as an overwhelming system constraint requiring institutional restructuring rather than technological solutions.

The Occupational English trainer (OET) managed 32 working professionals with 5–8 hours/week grading, characterizing delayed feedback as "lower priority compared to access and practice barriers." She noted: "My students are focused on practical competence, not grades...Most learning happens through repetition and practice, not through written feedback." For her student population, feedback timing was secondary to opportunities for repeated workplace-relevant communication practice.

These differences were not idiosyncratic preferences but systematically related to class size, workload, and professional context. The proficiency-differentiated impact was consistent across all three trainers: lower-proficiency learners internalized repeated mistakes when feedback was delayed and seemed less able to apply delayed feedback to future work; higher-proficiency learners adapted better to delayed feedback but still benefited from faster turnaround enabling more advanced guidance.

4.1.2 Attitudes Toward AI-Enabled Self-Assessment: Conditional Pragmatism

All three trainers acknowledged AI's potential for providing immediate feedback on surface errors while expressing legitimate concerns about learning depth, equity barriers, and enforcement challenges. The Business trainer stated: "I have mixed feelings, honestly. On one hand, if students could use a tool like Grammarly to catch their own grammar mistakes before submitting, that would help...I could focus on more important issues. I worry about dependence. If students rely on AI tools to correct their writing, do they actually learn the grammar rules?" The Academic trainer expressed similar pragmatic openness: "On the positive side, if students could catch their own grammar and organization issues before submitting, that would reduce my marking load. On the negative side, I'm concerned about academic integrity and depth of learning." She estimated realistic workload reduction at 20–30%, not the 40–50% sometimes claimed. The Occupational trainer was most skeptical: "If a tool could help students catch their own mistakes, that's efficient. On the other hand, my students often lack basic English skills. They don't have the foundation to independently evaluate whether an AI suggestion is correct." Her concern centered on foundational competence development: "I worry that students will accept AI corrections without understanding why they're correct, which doesn't build learning."

Critically, all trainers explicitly distinguished AI-for-checking from AI-for-writing. The Business trainer stated: "Using AI to check grammar is like using a spell-checker or a dictionary,it's a tool for refinement. Using AI to write the ideas is outsourcing the thinking, which is not learning." This distinction was conceptually clear but acknowledged as practically ambiguous in classroom implementation, requiring explicit teaching and enforcement.Viability assessments were not abstract judgments but practical calculations based on: (1) whether trainers experienced delayed feedback as their primary barrier; (2) whether they had workload capacity to implement new solutions; (3) whether student populations had prerequisite skills; (4) whether institutional conditions supported implementation.

The Business trainer assessed AI self-assessment as "genuinely viable" if clear distinctions between checking and writing were taught, enforcement mechanisms existed, and institutional policies aligned. She estimated realistic workload reduction: "Potentially yes...I might reduce my grading time from 10–12 hours per week to 7–8 hours per week."The Academic trainer assessed viability as "theoretically possible but practically limited" due to overwhelming workload: "It could reduce my marking time, maybe 20–30%...But I'm still spending weeks waiting for feedback." Without workload reduction, she lacked capacity to implement thoughtfully.The Occupational trainer assessed viability as "not realistic for my context" due to equity barriers (students lack device access) and foundational skill gaps: "Not really. It assumes students have access to tools many don't have. It assumes delayed feedback is the primary barrier, which it isn't in my context."

4.2 Research Question 2: Learner and Trainer Attitudes Toward AI-Mediated Assessment

4.2.1 Student Attitudes Toward AI-Assisted Self-Assessment

Survey data from 22 ESP students (M = 3.64/5.0, SD = .82 across self-assessment section) revealed moderate-to-strong enthusiasm for AI-assisted immediate feedback. Students reported highest agreement (M = 4.23/5.0) with "willingness to use AI tools for immediate self-assessment feedback," and strong agreement with the proposition that "using AI to fix surface errors differs fundamentally from having it write one's ideas" (M = 4.05/5.0). This 0.46-point difference was significant: students possessed sophisticated understanding of AI's appropriate role.

Qualitative responses provided granular insight into this nuanced perspective. Eighteen of 22 students explicitly stated they would not allow AI to write complete answers, and twelve emphasized opposition to complete replacement of student work. One respondent articulated a comprehensive framework:

"What I allow: using AI tools to simplify grammar rules and request examples to understand them better, or explain different types of errors and clarify their cause after the student completes their work...What I don't allow: using it to write answers instead of the student, copying answers as they are and submitting them without adding the student's touch...Reason: Because these rules ensure that AI is a helping tool for the student, not a substitute for the student's effort" (Student 12).

Students reported moderate frustration with delayed feedback (M = 3.50/5.0) and expressed specific learning needs unmet by current systems. Eight students mentioned "not knowing if work is correct" as a significant concern, while seven noted difficulty interpreting feedback without immediate opportunity for clarification. This uncertainty gap appeared directly addressable through immediate AI feedback mechanisms paired with instructor guidance.

4.2.2 Academic Integrity and Skill Development Concerns

Despite enthusiasm for AI assistance, students exhibited legitimate concerns (M = 3.73/5.0) about whether using AI might be considered cheating without clear rules, and uncertainty about which uses are acceptable (M = 3.41/5.0). These concerns were not hypothetical: 41% of respondents reported their current courses lack clear AI policies.Students also acknowledged fear that relying on AI might weaken their own skills (M = 3.50/5.0). Qualitative responses elaborated on this concern. One student wrote: "I will not allow writing what was asked of them instead of them because it leads to laziness, weak linguistic resources, and reliance on quick methods instead of trying." Another emphasized: "Learning and acquiring a language is not linked to AI tools, it is linked to the person themselves. I do not advise them to rely on it completely because it loses confidence and lack of initiative." However, students simultaneously acknowledged that these concerns differed from categorical opposition to AI use. One respondent provided particularly thoughtful reflection: "I want to emphasize that human feedback cannot be replaced no matter how advanced AI becomes. AI is like a mirror, not a crutch." This metaphor captured the desired relationship: a tool for reflection and self-assessment rather than dependency.

Trainer perspectives aligned with student concerns about skill development, though their emphasis varied by context. The Business trainer focused on understanding rules: "Students need to understand the difference between checking and writing...evaluate AI feedback critically." The Occupational trainer emphasized foundational competence: "My students need to develop confidence and independence...I'm not sure AI tools help with that." The Academic trainer highlighted deep critical thinking: "For substantive academic writing feedback 'Your argument is underdeveloped' I'm not sure AI can match human judgment."These concerns were highest in occupational and academic contexts where foundational competence or deep critical thinking mattered most, and lowest in business contexts where rule-following is more teachable and trainer-led instruction can scaffold AI feedback effectively. All trainers explicitly recognized critical boundaries between appropriate and inappropriate AI use, and all students demonstrated this distinction in qualitative responses. However, boundaries proved conceptually clear but practically ambiguous in implementation.

4.2.3 Boundary Recognition and Enforcement Concerns

The Business trainer noted: "The challenge is that I cannot easily tell the difference between a polished student email and an AI-generated email." He estimated that "in any group of 30 students, 5–10% would be tempted to use AI beyond just checking, especially under pressure before a deadline.".The Academic trainer acknowledged: "If a weaker student's essay is suddenly well-organized and clearly argued, that's suspicious...A student could give ChatGPT their essay topic and parameters, and ChatGPT would generate a reasonable academic essay." But she also acknowledged detection difficulties: "Stylistic inconsistencies might suggest AI, but I might not always detect AI use reliably." Critically, all trainers agreed that academic integrity concerns alone should not drive rejection of AI-mediated assessment. Rather, concerns should shape implementation design including verification mechanisms (in-class writing, process documentation, explanatory interviews) beyond rule-enforcement. The consensus position was clear: "rules are necessary but insufficient; institutional frameworks with multiple verification approaches are essential."

4.3 Research Question 3: Attitudes Toward Co-Construction

4.3.1 Trainer Receptivity: Contextually Determined

Co-construction receptivity exhibited significant contextual variation directly related to structural conditions.

Business English Trainer (Willing to Try). Position: "Yes, I would be willing to try it. I see potential value in it...My initial reaction is cautiously interested." Reasoning: He experienced moderate workload permitting experimentation, genuine feedback delay problem she wanted to solve, and saw pedagogical value in "teaching students to think critically about technology." She noted alignment with business education (business students study organizational policies and ethics). Requirements: "Clear guidance on how to facilitate the discussion," "departmental support," "clarity on enforcement," and "time allocation recognition." Assessment: "It's a genuine priority, actually."

Academic English Trainer (Skeptical Due to Capacity). Position: "Reluctantly, maybe...I'm too overloaded to take this on well...I'm not dismissing the idea. I'm saying my current workload makes it hard to imagine taking this on." Reasoning: Severe workload (15–20 hrs/week, 120 students) made adding another initiative impossible without removing something else. Not philosophical opposition but practical impossibility. Alternative position: "If I had fewer students, more resources, institutional support, I would be interested." Assessment: Her skepticism was capacity-constrained, not principled.

Occupational English Trainer (Resistant on Equity Grounds). Position: "No, I would not prioritize this...My reaction is skeptical." Reasoning: Saw co-construction as disconnected from real student needs, most lack reliable technology access; many lack relevant experience with AI tools; and delayed feedback was not her biggest problem. Equity focus: "Students with home internet would use AI; others wouldn't." Preferred institutional investment first: "If the proposal were about getting better technology infrastructure..." Assessment: Resistance was principled and equity-focused, not dismissive of participatory approaches generally.

Critical insight are drawn to indicate that the variation in receptivity was not random. Structural conditions determined attitudes;Business trainer had conditions enabling willingness, Academic trainer lacked capacity, Occupational trainer faced equity barriers.

4.3.2 Student Attitudes Toward Co-Construction

Students expressed strong support for participatory policy development (overall M = 3.64/5.0, with 4.05/5.0 for "wanting voice in defining boundaries"). This finding had important implications: students were not merely asking for permission to use AI but requesting voice in establishing appropriate use boundaries. Support was similarly high for related propositions: students would feel safer using AI if rules were co-created (M = 3.77/5.0), would be more committed to following co-created rules (M = 3.77/5.0), and believed co-construction would clarify appropriate usage (M = 3.59/5.0).The qualitative responses reinforced this finding. One student wrote: "I will trust rules that we created together more than rules imposed from above." Another stated: "If we help make the rules, we'll understand why they exist, not just follow blindly." This alignment with procedural justice literature indicated that participation increased perceived legitimacy and compliance motivation.

4.3.3 Comparison With Bans and Instructor-Only Approaches

All trainers universally rejected bans as unenforceable and counterproductive. The Business trainer stated: "A ban would be unenforceable anyway,students are already using AI tools...I would lose the opportunity to guide their use." The Academic trainer similarly noted: "A ban would be difficult to enforce and hypocritical...Students would use tools secretly, and I would lose ability to guide their use."Instructor-only policies were acknowledged as inadequate by all trainers because they: created student confusion about boundaries; lacked institutional consistency; didn't build genuine understanding; failed to address underlying problems. The Business trainer admitted: "I haven't created enough clarity, which is partly my responsibility" with her current informal syllabus note.

The comparison revealed recognition that effective approaches required somewhere between "no guidance" (bans) and "top-down mandates" (instructor-only). Recommended alternatives included: (1) structured trainer-developed guidelines with student input/feedback; (2) differentiated policies by class/discipline; (3) graduated engagement starting with consultation, progressing toward co-construction as comfort increased.

4.4 What Would Need to Change for Co-Construction to Succeed

Trainers identified specific, addressable barriers.Business Trainer claims that **,**for Co-Construction to Succeed , it needs "clear guidance on how to facilitate," "departmental support," "clear process for enforcement," and "time allocation recognition." Assessment: "It's ready to pilot now." For Academic Trainer it simply needs "workload reduction" (one fewer course or smaller classes), "structured guidance," "institutional backing," and "technology support." Assessment: "I'm too overloaded to take this on well" without systemic changes.Occupational Trainer,on the other hand , states ir requires "technology infrastructure (computers, reliable internet, institutional AI access)," "professional development on AI literacy," "clear institutional policy framework," and "equity considerations." Assessment: "This is a multi-year, institution-wide project."

On may critically interpret that the barriers to co-construction were not philosophical but institutional and structural(workload, technology access, equity, and institutional support). These barriers were addressable through hiring, infrastructure investment, professional development, and policy alignment.

5. Discussion

5.1 Integrated Findings: Context Shapes Receptivity

The most critical finding was that trainer attitudes toward delayed feedback, AI viability, and co-construction were not idiosyncratic preferences but systematically tied to structural conditions. The three trainers represented three different institutional realities: The Business trainer, experiencing moderate workload and concrete feedback delay problem, demonstrated pragmatic openness to both AI-assisted self-assessment and co-construction. The Academic trainer, experiencing systemic resource crisis, expressed capacity-constrained skepticism that could shift toward openness with workload reduction. The Occupational trainer, facing equity barriers rather than feedback delays as primary constraint, demonstrated principled resistance that reflected genuine concern for equitable access rather than opposition to innovation.This variation offers important theoretical insight.The trainer attitudes are not fixed traits but contextually-determined responses to structural conditions. To change receptivity, institutions must change the structural conditions enabling or constraining action.

The research confirms that AI-mediated feedback addresses real problems (surface-level errors, immediacy) while exhibiting real limitations (cannot replace human judgment for complex feedback, depends on student foundational skills, requires guidance to prevent overdependence). The concept of "human-AI orchestration" emerged as pedagogically sound: AI handles surface-level, quick-turnaround feedback; humans provide strategic, contextual, developmental feedback that automated systems cannot replicate (Taylor, 2024).

Effective implementation requires explicit teaching of when, why, and how to use AI appropriately;not assumption that students automatically develop critical engagement capacities. Shi et al. (2025) found that critical evaluation of AI-provided information remained lower than desired, highlighting the need for additional scaffolding. The research supports this finding: while students demonstrated sophisticated conceptual understanding of boundaries, implementation would require ongoing guidance.

5.2 AI as Partial Solution Within Broader Pedagogy

Co-constructed policy development proved educationally sound: participation in deliberation about appropriate AI usage develops AI literacy through authentic engagement with critical technology questions. The concept of "Students as AI Literate Designers," emphasizing that design cycle participation improves AI skills and ethical awareness, receives support from these findings (Shi et al., 2025). However, co-construction proved institutionally demanding in resource-constrained contexts.

The research revealed that co-construction should be incentivized and resourced rather than mandated universally. Trainers willing and able (those with moderate workload and clear problems to solve) should receive support and celebration. Those lacking capacity should not be pressured; alternative participatory approaches (consultation, input-gathering) should be acceptable for under-resourced contexts. Institutional policies should create enabling conditions through workload restructuring, infrastructure investment, and professional development support.

5.3 Limitations and Implications

The study examined three trainers and 22 students from two institutions , limiting generalizability. Findings should be interpreted as qualitative insights generating hypotheses rather than statistically representative claims. Additionally, cross-sectional design captures attitudes at a single moment; longitudinal investigation of how attitudes shift after exposure to co-construction would strengthen evidence about receptivity.

However, the contextual analysis of how structural conditions shape attitudes offers important insights for practice. The finding that institutional conditions (workload, technology infrastructure, equity) fundamentally constrain or enable receptivity suggests that policy reform requires systemic rather than individual-level interventions.

6. Implications and Recommendations

Institutional effectiveness requires multifaceted strategies addressing implications at both policy and operational levels. At the policy level, institutions should adopt differentiated rather than unified approaches, developing overarching frameworks permitting contextual adaptation across different ESP specializations, class sizes, and student populations, using tiered models such as the EAP-AIAS framework to guide appropriate AI integration parameters for different assessment types (Roe, 2024). Before launching co-construction initiatives, institutions must invest in foundational conditions that enable meaningful participation: reducing class sizes or providing teaching assistants to address trainer workload constraints, investing in technology infrastructure to ensure equitable device access, and providing professional development in participatory facilitation techniques that help educators navigate power dynamics and translate stakeholder input into implementable policies. Institutions should position co-construction as exemplary rather than mandatory practice, creating incentive structures and resources specifically for trainers willing to engage in participatory policy development while recognizing and celebrating successful initiatives; for under-resourced contexts lacking capacity for intensive co-construction, institutions should accept alternative participatory approaches—such as structured consultation, feedback-gathering, or graduated engagement starting with consultation and progressing toward co-construction as comfort increases as legitimate and equally valuable contributions to policy development (Partnership on AI, 2023). Institutionally, providing templates, frameworks, and facilitation guidance for conducting policy co-construction discussions ensures consistency and quality across initiatives, including training on navigating power dynamics, ensuring genuine rather than nominal participation, and translating diverse stakeholder perspectives into coherent, enforceable policies (Chaudhry et al., 2024).

Trainer professional development must parallel institutional policy reform, with institutions supporting educators in developing comprehensive AI literacy encompassing not merely technical capabilities but pedagogical roles and limitations—explicitly including discussion of human-AI orchestration as framework for thoughtful integration rather than either blanket prohibition or wholesale adoption without safeguards (Taylor, 2024). Critically, professional development initiatives must acknowledge that trainer capacity represents a fundamental prerequisite for meaningful implementation; without addressing workload constraints that limit educator availability and cognitive bandwidth, even well-designed professional development programs will fail to achieve adoption. Institutions must advocate forcefully for hiring investments and structural reforms reducing grading burden, which enables genuine engagement with new initiatives rather than adding technological solutions to already overwhelming workloads.

Future research directions should prioritize mixed-methods longitudinal studies tracking how co-construction processes unfold over time and impact not merely immediate attitudes but sustained learner outcomes, trainer practices, and institutional cultures, strengthening the empirical evidence base for policy recommendations (Wilson et al., 2025; Fuligni et al., 2025). Comparative studies examining co-construction across different ESP specializations, institutional contexts, and cultural settings would illuminate which contextual factors most significantly shape implementation success and which elements can transfer across contexts. Research specifically examining how institutional restructuring initiatives(workload reduction, infrastructure investment, professional development support) enable or constrain the feasibility of co-construction would provide essential evidence directly applicable to institutional planning and advocacy (Nadarzynski et al., 2025).

7. Conclusion

The integration of artificial intelligence into ESP education offers transformative potential for addressing persistent feedback challenges(delayed turnaround and insufficient disciplinary specificity) that constrain learning effectiveness. However, realizing this potential requires governance structures that honor learner agency, respect trainer expertise, and acknowledge that optimal policies emerge through sustained dialogue among affected stakeholders rather than unilateral mandate.

This feasibility and needs analysis study examined whether co-constructed AI policies(collaboratively developed by learners and trainers) constitute a viable alternative to restrictive bans or instructor-only approaches. Findings reveal significant contextual variation in how stakeholders experience AI as solution and view participatory processes as appropriate governance mechanism. Business English trainers with manageable workloads and concrete feedback delay problems demonstrate pragmatic openness. Academic English trainers with overwhelming workloads express capacity-constrained skepticism. Occupational English trainers with equity barriers demonstrate principled concerns. Rather than random variation, this differentiation reflects logical consequence of structural conditions shaping whether individuals have capacity and motivation to engage in co-construction.

ESP students, by contrast, demonstrate sophisticated understanding of appropriate AI boundaries, strong preference for co-constructed rules over bans or instructor-only policies, and confidence in their capability for participatory engagement. Their voices reveal not naïveté but nuanced understanding, not resistance to standards but desire for clarity, and not rejection of teacher expertise but request for collaborative partnership.

Co-construction represents both aspirational ideal and potential practice,but only when institutions provide foundational conditions enabling genuine participation: manageable trainer workload, equitable technology access, professional development in facilitation, and institutional policy frameworks supporting rather than impeding collaborative decision-making. The research contributes to educational technology governance scholarship by demonstrating that effective AI integration cannot be imposed from above but must emerge through collective wisdom of those directly affected by policies. For ESP contexts confronting persistent feedback challenges, this approach offers mechanism to improve learning immediacy while developing critical technology literacy and honoring educational democracy,not despite stakeholder involvement in decision-making, but precisely because of it.

References

Alshehri, A. (2025). AI-powered personalized feedback and EFL learner outcomes: A mixed-methods investigation. *Language Learning & Technology, 29*(1), 1–22.

Assassi, N. (2025). Leveraging generative AI in English for Specific Purposes instruction: Opportunities and challenges. *Computer Assisted Language Learning, 38*(2), 1–18.

Chaudhry, A. S., Zhao, Y., & Liu, H. (2024). Transparency requirements for AI-based educational technology tools: A stakeholder engagement approach. *AI and Ethics, 4*(2), 1–15.

Dekhakhena, K. (2025). Real-time AI feedback and speaking fluency development: Evidence from intermediate EFL learners. *System, 98*, Article 103054.

Fuligni, A. J., Chen, L., & Wilson, R. (2025). The Contextualized Perceptions for the Adoption of LLMs in Education (Co-PALE) framework: Prioritizing stakeholder perspectives in AI governance. *Technology in Society, 45*, Article 102340.

Huang, X., Li, Z., Chen, J., & Park, K. (2023). AI affordances in language learning: A systematic review and research agenda. *Educational Technology Research and Development, 71*, 945–968.

Huang, X., Li, Z., Zhu, C., Wang, Y., & Gao, H. (2024). ChatGPT adoption in higher education: Insights from a comprehensive analysis of 500 universities worldwide. *Computers and Education: Artificial Intelligence, 5*, Article 100153.

Huang, X., Wang, Y., & Lee, S. (2024). Integrating ChatGPT in English for Specific Purposes pedagogy: Opportunities and implementation challenges. *TESOL Quarterly, 58*(1), 45–67.

Liebenow, C., Murphy, K., & Thompson, B. (2025). LLM-generated feedback and student self-assessment accuracy: When does AI help? *Computers & Education, 198*, Article 105121.

Moorhouse, B. L., & Kohnke, L. (2024). AI in language teaching and learning: Current state, challenges, and emerging directions. *TESOL Quarterly, 58*(2), 234–256.

Nadarzynski, T., Narenji Sheshgelani, A., Wang, W., & Holmes, S. (2025). Participatory design of chatbot-assisted self-assessment: Integrating stakeholder perspectives. *International Journal for Participatory and Experiential Research, 8*(1), 1–24.

Partnership on AI. (2023). *Guidelines for participatory and inclusive AI*. Partnership on AI.

Plastina, A. F. (2015). Learner engagement and feedback perception in blended learning: A case study of English for Medical Purposes. *Journal of English for Academic Purposes, 19*, 33–41.

Roe, C. L. (2024). The EAP-AIAS: Developing an assessment scale for GenAI integration in English for Academic Purposes. *English Teaching Forum, 62*(3), 2–10.

Shi, J., Liu, C., & Wang, Y. (2025). Exploring how AI literacy and self-regulated learning relate to writing performance and digital well-being. *Humanities and Social Sciences Communications, 12*, Article 985.

Stanišić, N., Maksimović, J., & Osmanović, J. (2022). Implementation challenges of ESP blended learning in higher education contexts. *International Journal of Education, 16*(2), 117–132.

Taylor, L. (2024). Blended learning challenges of EFL undergraduate students: Student learning experience in an AI-integrated ESP course. *Studies in English Language and Education, 11*(1), 123–138.

Wilson, J., et al. (2025). Towards child-centred AI in children's learning: The Co-Design Stories Toolkit. *International Journal of Human-Computer Studies, 145*, Article 103245.