Theory and Practice of Second Language Acquisition
vol. 12 (2), 2026, pp. 1/33
https://doi.org/10.31261/TAPSLA.21925

CC BY SA

Robert Oliwa
University of Rzeszów, Poland
https://orcid.org/0000-0002-5251-1086

Integrating AI-Generated Avatars into Nursing ESP: Student Perceptions, Readiness and Implications for Professional Communication

Abstract

This study investigates the use of AI-generated avatars in English for Specific Purposes (ESP) instruction for nursing students, focusing on their perceptions, preferences and readiness to adopt avatar-supported learning and examining how these factors influence professional English communication in healthcare contexts. A quasi-experimental design compared two groups: an experimental group, where avatars were integrated into the ESP syllabus under teacher guidance, and a control group, which followed the traditional coursebook with only self-directed access to the avatar. Findings show that the experimental group developed increasingly positive perceptions of avatar use, stronger preferences for its integration and greater readiness to adopt the technology. Teacher mediation was critical, aligning avatar activities with coursebook content and situating them within realistic scenarios such as patient handovers and physical assessments. By contrast, the control group, without structured integration, engaged less consistently with the avatar and reported weaker perceptions of preparedness for authentic clinical communication.

The results indicate that avatars enhance readiness through scenario-based practice and low-stakes rehearsal with immediate feedback, building both competence and confidence. However, concerns about limited human warmth and ethical issues such as privacy and equity highlight the need for balanced, responsible implementation.

Keywords: AI, avatar-supported learning, nursing, ESP

The accelerating pace of globalisation and technological advancement has redefined the skill sets required in the global workforce emphasising language proficiency, cultural competence and digital fluency. Artificial Intelligence (AI) is emerging as a critical factor shaping how languages are taught, learned and applied in diverse professional contexts. Incorporating AI into language education promises substantial potential in bridging traditional pedagogical gaps

and providing immersive, personalised and context-specific learning experiences. However, as AI increasingly permeates educational landscapes, there is an urgent need to systematically evaluate its efficacy, applicability and ethical implications within the context of language learning for global careers.

Despite the proven advantages and transformative potential of AI in language education, several critical issues remain unresolved. These encompass balancing personalised AI-supported learning experiences and necessary human interaction, concerns around data privacy and ethical usage, equity in AI accessibility as well as the potential risk of excessive reliance on technology potentially undermining foundational interpersonal skills (Pegrum, 2026; Roe, Perkins, & Furze, 2025). Additionally, while initial results indicate improved linguistic competence and decreased learner anxiety in professional contexts, comprehensive research examining the efficacy and practical integration of AI-driven virtual avatars into real-world educational settings remains scarce (Alsager, 2024).

There is a pressing need to rigorously assess how effectively AI technologies, especially virtual avatars simulating professional roles, enhance language learning outcomes and adequately prepare students for the linguistic demands of global careers. Critical investigation is required into students' perceptions, preferences and readiness to accept avatar-supported learning as it becomes widely used. In this study, readiness is understood as students' perceived preparedness and willingness to engage with AI-supported English language learning, particularly through the use of an AI-driven avatar. It is approached as a composite of attitudes toward AI, perceived relevance for professional language tasks, ethical comfort and intended use, rather than as a purely technical competence. Readiness is therefore treated as context-dependent and experience-sensitive, shaped by prior digital exposure and, where applicable, direct interaction with AI-supported learning environments.

This research systematically attempts to investigate these dimensions to offer actionable insights for leveraging AI technologies effectively, ethically and equitably in global language education, directly informing educators, policymakers, curriculum designers and learners preparing for global professional environments.

Review of Literature

Artificial intelligence has become increasingly embedded in language education, yet its pedagogical value remains contingent on how it is implemented rather than on technological sophistication alone. In professional language

contexts, the central question is not whether AI can generate language, but whether it meaningfully supports communicative development aligned with domain-specific needs.

AI literacy—understood as the capacity to understand, use and critically evaluate AI tools—has been associated with learners' willingness to communicate and perceived competence (Zhang et al., 2025). Enhanced AI-related self-efficacy may influence engagement in AI-supported environments. However, literacy alone does not guarantee pedagogical effectiveness. Pegrum (2026) argues that AI integration must remain anchored in ethical reasoning and human-centred instructional design. Similarly, Liu (2025) emphasises that educators require not only technical familiarity but the ability to align AI use with curricular aims and professional learning outcomes. In language education, teacher mediation therefore remains central; AI systems may supplement instruction, but they do not replace pedagogical judgment (Pegrum, 2026).

Recent developments—including adaptive systems, speech recognition and intelligent tutoring technologies—have expanded possibilities for context-sensitive feedback (Zhu & Wang, 2025). Of particular relevance to ESP are AI-driven avatars capable of simulating professional interaction. Such systems can generate scenario-based dialogue and provide targeted feedback, potentially supporting communicative rehearsal in domain-specific contexts (Peña-Acuña & Corga Fernandes Durão, 2024; Alsager, 2024). While generative AI tools such as ChatGPT are widely used for text-based support, their function differs from interactive avatars, which are designed to simulate situated communication rather than assist with discrete language tasks.

Healthcare communication presents particular demands. Nurses and other healthcare professionals must obtain accurate patient histories, explain procedures, deliver instructions and respond to emotional cues in culturally diverse settings. ESP theory underscores the importance of aligning instruction with authentic professional tasks (Hutchinson & Waters, 1987; Dudley-Evans & St John, 1998). Scenario-based simulation reflects task-based principles (Ellis, 2003) by situating language within purposeful action rather than isolated form practice. Within this framework, AI-driven avatars may serve as structured rehearsal environments, approximating patient interaction while remaining pedagogically controllable (Alsager, 2024).

Computer-assisted language learning research similarly highlights the importance of meaningful input, opportunities for output and timely feedback (Chapelle, 2001). Adaptive systems, including intelligent tutoring applications (Heift & Schulze, 2007), may strengthen form–meaning connections when integrated within coherent instructional sequences. Lower-stakes rehearsal environments have also been associated with increased willingness to communicate, although affective effects remain context-dependent (Krashen, 1982). These

theoretical strands converge in suggesting that contextualised interaction, feedback and alignment with professional tasks are critical conditions for learning.

Ethical and structural considerations shape how AI can be responsibly integrated. Concerns regarding data privacy, algorithmic transparency and equitable access have been widely noted (Chiu et al., 2023). In healthcare education, where simulated scenarios mirror ethically sensitive professional contexts, such considerations carry additional weight. Pegrum (2026) cautions against treating AI as a universal solution, particularly where relational and interpersonal dimensions of learning are central. Peña-Acuña and Durão (2024) similarly stress the need for institutional responsibility in ensuring equitable access and pedagogical coherence. While digital divides are not unique to AI, uneven implementation may exacerbate disparities if institutional support is absent.

Existing literature has tended to focus either on general AI affordances or on isolated tool usage rather than on structured curricular integration within specific professional domains. Pegrum (2026) warns that generalised AI applications may inadequately address specialised communicative demands. Alsager (2024) and Peña-Acuña and Durão (2024) note the absence of sustained empirical evidence demonstrating long-term impact on professional readiness. Chiu et al. (2023) call for context-specific AI applications grounded in authentic task design and evaluated through rigorous pedagogical criteria.

These perspectives suggest a potential rationale for the use of AI-generated avatars in ESP contexts. From an SLA standpoint, language development is closely associated with opportunities for interaction, particularly where learners are required to negotiate meaning and produce extended output. In this sense, avatars may function as accessible interlocutors, creating conditions in which such interaction can occur more frequently than in conventional classroom settings. At the same time, ESP frameworks emphasise the importance of situated learning, where language use is embedded in domain-specific tasks. When avatars are integrated into clinically oriented scenarios, they may provide a context in which linguistic forms are more directly linked to professional practice.

From a CALL perspective, these processes may be further shaped by the availability of immediate and repeatable feedback, as well as by the possibility of sustained, self-paced engagement. Feedback of this kind may support noticing and gradual refinement, while repeated interaction may allow learners to revisit and adjust their responses over time. However, these affordances are unlikely to operate uniformly across contexts. Their effectiveness may depend on how the tool is embedded within instruction, particularly whether interaction is structured and supported or left largely self-directed.

In this light, AI-generated avatars can be understood as environments in which interaction, situated practice and feedback potentially converge. This provides a basis for examining whether different modes of integration—such

as guided versus unstructured use—are associated with distinct patterns in students' readiness and communicative development.

Research Methodology

Research Objectives

The study set out to examine how the integration of an AI-generated avatar within a nursing ESP course relates to students' perceptions, readiness and communicative development. To address this aim, three more specific objectives were formulated.

First, the study compares the perceptions of the experimental group (EG) and the control group (CG) regarding the role of AI in language learning. This comparison draws on questionnaire data collected at baseline and post-intervention and is analysed through changes in response distributions, with particular attention to shifts in agreement and uncertainty. Second, the study examines changes in students' readiness to adopt AI-supported learning over time. Readiness is operationalised through questionnaire items addressing willingness to participate, intended frequency of use and perceived usefulness of the avatar. These indicators are analysed descriptively by tracing directional changes between the two measurement points. Third, the study evaluates the role of teacher-mediated integration in shaping engagement and communicative development. This objective is addressed through a comparative interpretation of (a) structured avatar use in the EG, (b) self-directed use in the CG and (c) performance-based measures, including classroom assessments, Cambridge placement results and avatar-generated feedback on spoken interaction.

The above objectives allow for a differentiated analysis of how structured versus unstructured exposure to AI-supported learning relates to attitudinal change and context-specific communicative performance.

Research Participants

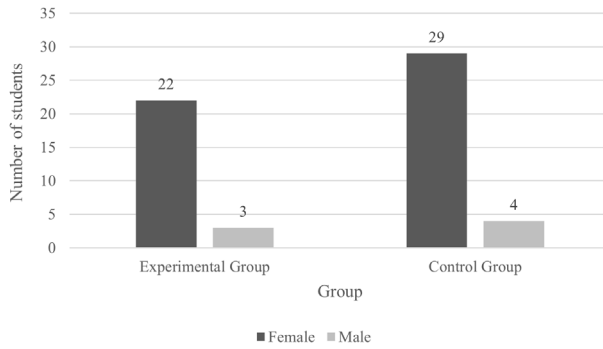
The research was conducted among two distinct groups of first year nursing students. Participants had a comparable background in general English education and no prior exposure to ESP in healthcare contexts. Several students reported having studied German as part of their school curriculum, while others indicated additional language learning through private courses, most commonly in English and occasionally in Spanish. One participant reported professional

experience in Italy and described functional spoken competence in Italian. The students were enrolled in a higher education nursing programme, with no prior ESP or professional healthcare experience.

The selection of participants was purposeful, focusing on learners in healthcare-related education, where communicative competence in English is particularly consequential in both professional and international contexts. In total, the sample comprised 58 students, including 51 women and seven men, reflecting typical gender distributions in nursing education. The experimental group (**EG**) consisted of 25 students (22 women and three men), while the control group (**CG**) included 33 students (29 women and four men) (see Figure 1).

Figure 1

Group Distribution by Gender (EG vs CG)



Source: Author's research.

Both groups, as part of their regular English language classes, follow the coursebook *Career Paths: Nursing* by Express Publishing (Evans et al., 2019). This specialised ESP coursebook is designed for nursing students and professionals, offering a structured and progressive curriculum across three levels of difficulty. It integrates the four key language skills—reading, listening, speaking and writing—into content specifically tailored to medical and clinical contexts. Each unit is built around a specific theme relevant to the nursing profession (e.g., hospital staff, patient symptoms, physical examination, administering medication or post-surgical care). This structure provided all participants with foundational exposure to workplace-relevant English, reinforcing lexical competence and communication strategies in professional settings.

The students of both researched groups were granted access to a fully prompted Soul Machines avatar (https://www.soulmachines.com) for language practice. Soul Machines is an advanced AI platform that creates highly realistic, lifelike digital humans capable of engaging in natural, emotionally responsive conversations. The AI avatar used in this study was a fully interactive digital

agent developed on the Soul Machines platform and configured as a professional English language coach for nursing students. It was assigned an explicit instructional identity: a professional, empathetic language instructor whose function was to guide and support learners rather than merely generate conversational output. In this sense, it operated as a quasi-teaching agent within a teacher-led curriculum.

Its instructional focus was English for nursing practice. The avatar supported the acquisition of medical terminology, simulated clinical communication (e.g., patient interviews and physical assessments) and provided feedback on pronunciation, fluency, grammatical accuracy and communicative appropriacy. Content was delivered through open-ended, real-time dialogue. Students engaged in scenario-based exchanges, during which the avatar responded contextually, adjusted linguistic complexity where necessary and offered immediate corrective feedback.

English functioned as the primary language of communication. Interaction design combined conversational AI with a defined behavioural profile (professional and encouraging tone) and recorded sessions in the EG for reflective follow-up. Importantly, the avatar did not determine curricular progression independently. In the EG, its use was structured and mediated by the instructor, who aligned tasks with the ESP syllabus and contextualised feedback.

Thus, while technologically autonomous in generating responses, the avatar remained pedagogically embedded. It modelled professional discourse, structured rehearsal of clinical scenarios and delivered automated formative feedback, yet curricular authority and interpretive framing were retained by the human teacher.

Despite the fact that both groups were granted access to the avatar, only the students of the EG interacted with the avatar in class under the direct supervision and guidance of the author. These sessions allowed for structured support, contextual clarification in addition to immediate pedagogical intervention when necessary. The author facilitated role-playing scenarios, provided scaffolding for difficult interactions and integrated follow-up discussions to reinforce key vocabulary and communicative strategies. In contrast, the students of the CG were sent the link to the avatar and encouraged to use it independently for self-study. While they had equal technical access, their use of the tool lacked the teacher's instructional support, real-time feedback from the teacher and collaborative classroom context.

The EG's experience was differentiated by the integration of the AI avatar, which extended and contextualised the coursebook's content into simulated, real-time clinical interactions. The avatar allowed students to practise role-playing scenarios (such as physical assessment) with personalised feedback, making their engagement more interactive and immediate. Meanwhile, the CG used the coursebook in its standard form, with only self-directed access to

the avatar, without the pedagogical mediation that helped align the tool with textbook content. Avatar usage was systematically tracked only in the EG, where all interactions took place during scheduled language classes under the author's supervision. Engagement was embedded in structured role-play activities integrated into the syllabus, allowing consistent monitoring of participation and frequency of use.

Although both groups were granted equal technical access to the AI avatar, the conditions under which it was used differed substantially. In the experimental group, engagement was structured, scheduled and embedded within classroom activities under teacher guidance. By contrast, the control group accessed the avatar on a voluntary basis, without formal integration into the syllabus and without systematic monitoring of frequency or depth of use. As a result, the extent of engagement in the control group cannot be determined with precision.

This difference should be considered when interpreting the findings. At the same time, it reflects a deliberate contrast rather than an unintended imbalance. The study was designed to distinguish between two modes of exposure: guided, pedagogically framed use and unstructured, self-directed access. In this sense, the variability observed in the control group is not treated solely as a limitation, but as part of the analytical context. It provides an approximate reference point for how the tool may be used in the absence of instructional support, thereby allowing the effects of structured integration to be interpreted more clearly.

In this way, the research explored how the coursebook's foundational ESP content, when supplemented by AI-generated avatars, could enhance but also influence students' perceived readiness for real-world clinical interaction as well as attitudes towards the AI component itself, assessing how the integration of a lifelike digital avatar impacted their engagement, confidence and acceptance of AI as a pedagogical tool within nursing education.

At the same time, much of the literature focuses on general-purpose tools or short-term use, offering limited insight into sustained, curriculum-based integration. This is especially evident in research on AI-generated avatars, which tends to emphasise technical affordances or immediate learner responses rather than their role in structured ESP instruction.

As a result, a gap persists between theoretical claims and context-specific evidence. There is still little empirical work examining how fully interactive avatars, embedded within a guided ESP framework, influence students' perceptions, readiness and professional communication. The present study addresses this issue within a nursing context.

These characteristics suggest a cohort with broadly comparable general English exposure, limited prior ESP experience and heterogeneous multilingual backgrounds, which provides an important context for interpreting both initial readiness levels and subsequent changes observed during the study.

Research Tools

Two custom-designed questionnaires (see Appendix 1 and Appendix 2) were administered digitally using Microsoft Forms, selected for its accessibility, cost-effectiveness and ease of implementation. Questionnaire 1 (baseline) consisted of 23 items organised into eight sections: (1) demographic and background information, (2) digital literacy and experience with technology, (3) English language proficiency, (4) perceptions of AI in learning English, (5) ethical considerations and data privacy, (6) technical access and inclusion, and (7) willingness to use AI avatars in education. Items in Section 4 were assessed using a five-point Likert scale (agreement-based) ranging from strongly disagree to strongly agree, a format chosen because it allows respondents to express graded attitudinal positions while remaining cognitively manageable for participants with varying levels of language proficiency. Additional questions (8) addressed preferred frequency of AI avatar use and willingness to participate in future AI pilot studies.

Questionnaire 2 (post-intervention) was a shorter version, omitting the first three sections and focusing specifically on students' perceptions, expectations, ethical concerns and readiness to adopt AI-supported learning. Items examining the perceived importance of specific AI avatar features were measured using a five-point Likert-type importance scale, ranging from definitely not important to definitely important. This scale was selected to capture relative value judgements rather than agreement per se, reflecting the evaluative nature of these items.

No formal pilot study was conducted prior to implementation. However, the instrument was designed with attention to contextual relevance, with items formulated to reflect students' actual learning environment (e.g., patient interaction, clinical communication, medical terminology), thereby anchoring responses in experience. The use of two measurement points (pre- and post-intervention) allowed for the observation of both change and a degree of internal consistency over time. Although formal reliability indices such as Cronbach's alpha were not calculated, due to the exploratory design and limited sample size, coherence was assessed through thematic grouping and the stability of response patterns across the dataset.

Data were collected anonymously via Microsoft Forms. Responses were voluntary and not linked to course assessment. Questionnaire data were interpreted alongside classroom assessments, an external placement test and avatar-generated feedback to support triangulation.

In addition to questionnaire data, both groups completed regular classroom assessments aligned with the syllabus. Classroom assessment grades (e.g., very good, good, satisfactory) were used as course-based performance indicators and are not directly equivalent to CEFR levels. While the Cambridge placement

test provides a standardised CEFR reference, classroom results reflect domain-specific ESP performance within the course context. Spoken performance in the experimental group was further supported by avatar-generated feedback during simulations. Assessments were graded using standardised rubrics applied consistently across both groups to ensure comparability. The avatar also provided automated feedback on pronunciation, fluency and communicative appropriacy during simulated interactions, supporting process-level observation.

Data analysis was limited to descriptive statistics (frequencies and percentages) to summarise students' responses and observable trends over time. To extend the descriptive account, a limited set of additional analytical procedures was considered in order to examine whether the differences observed between the EG and CG, as well as changes across the study period, might reflect more than incidental variation. The structure of the dataset, however, places certain constraints on the type of analysis that can be meaningfully applied. As the data derive largely from aggregated Likert-type responses and categorical distributions, the analysis remains oriented toward comparative tendencies rather than fully specified inferential modelling.

In practice, attention was directed to how response patterns evolved within and between groups. The experimental group, over time, appears to move more consistently towards agreement on several readiness-related items, while the control group shows a more uneven profile, with smaller shifts and a greater persistence of uncertainty. When viewed in distributional terms, these changes take on a clearer shape. Non-parametric comparisons—most notably chi-square-based examinations of response frequencies—were therefore used selectively to assess whether such shifts might plausibly be treated as systematic. In the case of the EG, the reduction in neutral responses alongside increased agreement suggests a degree of consolidation that is less evident in the CG. This tendency resonates with the interpretation developed later in the Discussion, where the role of structured, teacher-mediated engagement is foregrounded.

At the same time, the analytical scope remains deliberately cautious. The available data do not support robust parametric estimation and any attempt to derive precise effect sizes would rely on assumptions that are difficult to justify fully in this context. It is nevertheless possible to consider the direction and relative consistency of observed changes. From this perspective, the intervention may be associated with modest but discernible shifts in perceived readiness and communicative orientation, particularly within the experimental group.

These observations should not be overstated. The sample is relatively small, the design is quasi-experimental and the translation of ordinal responses into analytical categories introduces a degree of simplification. For these reasons, the statistical layer of the analysis is best understood as complementary. It does not establish causal claims, but it does lend additional support to patterns already visible in the descriptive findings.

Procedure

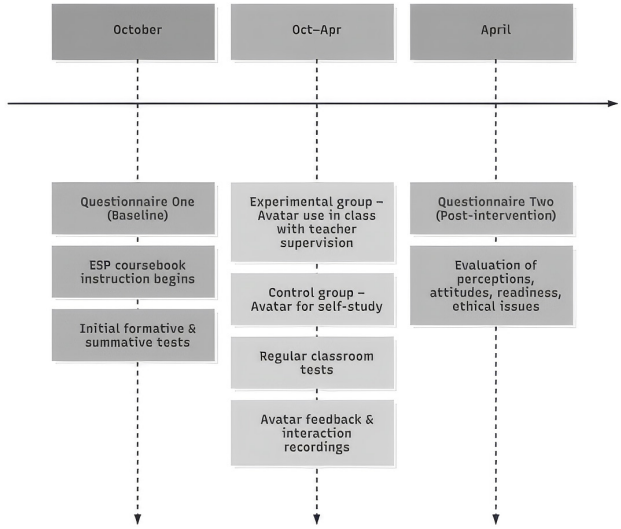
The study was conducted over the period from October to April and involved both an EG (guided classroom use under instructor supervision) and a CG (self-study with the avatar link). The procedure unfolded in three main phases (see Figure 2):

1. Baseline Phase (October)
   - At the start of the study, students in both groups completed Questionnaire 1, which gathered demographic data, information on digital literacy and English proficiency and baseline perceptions of AI in education.
   - Both groups also began their regular ESP coursework, following the *Career Paths: Nursing* syllabus, which included formative classroom tasks and scheduled summative tests.
2. Intervention Phase (October–April)
   - Throughout the period, the EG engaged with the AI avatar during class sessions, supported by the instructor. Interactions were scaffolded with follow-up discussions and integrated with coursebook content.
   - The CG was given access to the avatar link for independent self-study, without direct teacher involvement.
   - In both groups, regular formative and summative classroom tests were administered according to the coursebook schedule. These assessed students' acquisition of specialised vocabulary, grammar and reading and writing skills.
   - During all avatar sessions, the system provided immediate feedback on pronunciation, fluency and the accuracy of professional language use. Each session was automatically recorded, allowing students to revisit their interactions for reflection. For the EG, recordings were also used in class for additional teacher-led feedback and clarification.
3. Post-Intervention Phase (April)
   - At the end of the research period, students completed Questionnaire 2, designed to capture their experiences, perceptions and readiness to adopt avatar-supported learning after sustained exposure.
   - Questionnaire 2 focused specifically on AI-related perceptions, design expectations, ethical considerations and willingness to integrate AI into their future learning practices.

This procedure ensured that data were collected both before and after exposure to avatar-supported instruction, allowing for a comparative analysis of changes in attitudes and performance. The inclusion of classroom tests, avatar-generated feedback and recorded sessions created a triangulated dataset that combined subjective self-reports with objective measures of language learning outcomes.

Figure 2

Study Procedure Timeline (EG vs CG)



Source: Author's research.

Data Analysis

The analysis combines descriptive and comparative procedures to examine patterns across multiple data sources. Given the exploratory design and the structure of the dataset, emphasis is placed on identifying directional tendencies rather than on fully parameterised statistical modelling.

Questionnaire data are used to trace changes in students' perceptions, expectations and readiness between baseline and post-intervention stages. Responses, primarily collected through Likert-type and categorical formats, are analysed in terms of distributional shifts, with particular attention to changes in agreement, uncertainty and intended patterns of use. Where appropriate, these patterns are examined through indicative comparisons, including between-group contrasts (experimental vs control) and within-group changes over time. Selected distributions are also considered using non-parametric reasoning (e.g., chi-square-type comparisons), although the results are interpreted cautiously due to sample size and data structure.

Performance-related data provide a complementary perspective. Classroom assessments, aligned with the ESP syllabus, are treated as indicators of domain-specific linguistic development, while Cambridge placement results offer an external reference point for general language proficiency. In the experimental

group, avatar-generated feedback from clinical simulations adds a process-oriented dimension, capturing aspects of spoken interaction such as pronunciation, terminology use and communicative appropriacy.

These sources are interpreted in relation to one another. Converging patterns—for example, increased self-reported readiness alongside more consistent classroom performance—are treated as mutually reinforcing, whereas divergences are considered analytically informative, particularly where they reflect differences between general proficiency and context-specific competence. In this sense, the analysis remains integrative but not reductive, allowing attitudinal and performance-based data to inform a shared interpretation while retaining their distinct evidentiary roles. Overall, the sample is relatively small, the design is quasi-experimental and the transformation of ordinal data into analytical categories introduces a degree of approximation. Accordingly, the results are interpreted as indicative rather than conclusive, supporting the identification of emerging patterns rather than definitive causal claims.

Findings

In the analysis that follows, shifts in willingness, levels of uncertainty and preferred patterns of use are treated as observable indicators of readiness, consistent with its operationalisation outlined in the Research Tools section.

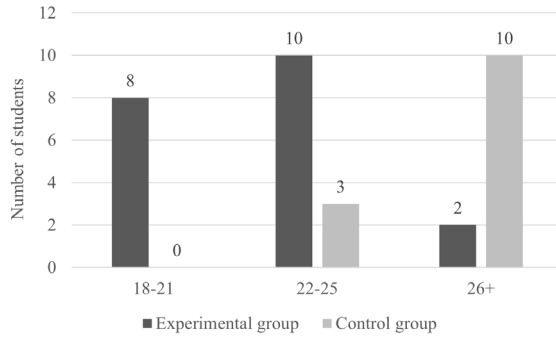
Age of the Participants

Questionnaire 1 was completed by 33 nursing students, representing 57% of the total cohort ($N = 58$). Although participation was voluntary, the responses provide a reasonable overview of the group's demographic profile. Most respondents were aged 22–25. Students aged 26 and above were more frequently represented in the control group, whereas those aged 18–21 appeared only in the experimental group. This uneven distribution should be considered when interpreting group comparisons, as age may have influenced learning preferences and engagement patterns.

In terms of age (see Figure 3), the experimental group (EG, N = 20) was predominantly young, with 18 students aged 18–25 and only two aged 26 or older. The control group (CG, N = 13) was considerably older: ten students were aged 26 or older, three were aged 22–25, and none were aged 18–21. With regard to gender, female students comprised the majority of respondents, which reflects common patterns in nursing education. At the same time, male participants were also represented, contributing a degree of gender diversity to the dataset.

Figure 3

Age Distribution by Group (EG vs CG)



Source: Author's research.

English Language Performance at the Start of the Research

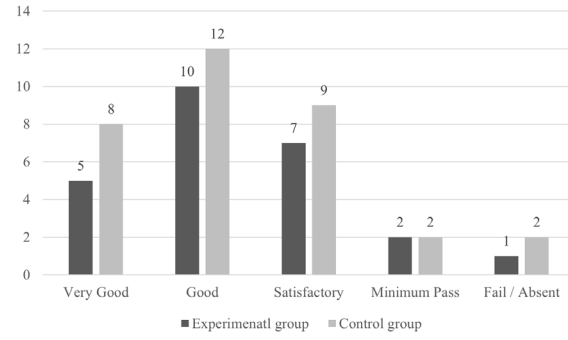
Baseline classroom assessments indicated broadly comparable proficiency levels across groups (see Figure 4), with some variation in grade distribution. In the EG, most students achieved good or satisfactory results, with a smaller number attaining very good grades. Two students were at the minimum pass level and one was absent or failed.

The CG showed a slightly stronger initial profile, with more students in the good and very good categories. However, as in the EG, a small number fell at the minimum pass level or were absent/failed.

Overall, while the CG included a somewhat higher proportion of stronger performers at baseline, both groups displayed mixed proficiency levels, including students near the lower performance threshold.

Figure 4

Baseline English Language Performance by Group (EG vs CG)



Source: Author's research.

English Language Performance at the End of the Research

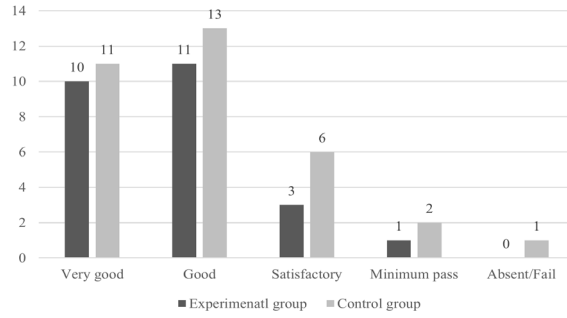
By the end of the research period in April, both groups showed progress in specialised vocabulary, grammar, reading and writing, as reflected in formative and summative assessments aligned with the syllabus (see Figure 5). In the EG, results clustered at the upper end of the grading scale, with most students achieving good or very good grades and no cases of failure.

The CG also suggested overall improvement, though performance was more dispersed. While many students achieved good or very good results, a larger proportion remained at the satisfactory level and a small number were assessed at the minimum pass level or recorded as absent/failed.

Overall, both groups met course objectives. The more concentrated performance profile in the EG may indicate greater consistency associated with guided integration, although this interpretation remains cautious given the descriptive nature of the analysis.

Figure 5

End-of-Study English Language Performance by Group (EG vs CG)



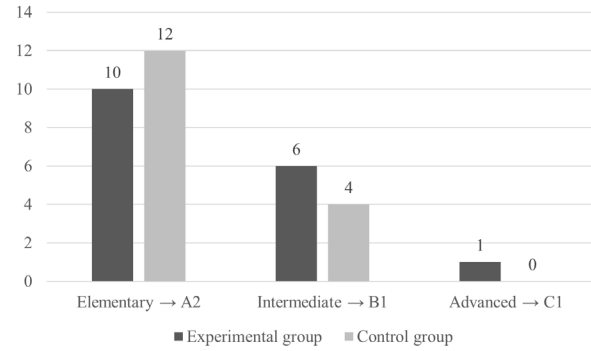
Source: Author's research.

English Language Proficiency

Participants' general English proficiency was additionally assessed using the Cambridge English online placement test, which served as an external benchmark alongside classroom evaluations. Results were concentrated at lower CEFR levels. Most students (22 of 33) were placed at A1, while ten reached B1. Only one participant (EG) attained C1 level (see Figure 6), indicating that advanced proficiency was rare within the cohort.

Figure 6

General English Proficiency Based on Cambridge Placement Test (EG and CG Combined)



Source: Author's research.

Some variation between classroom assessments and standardised test results can be observed. This may be partly explained by differences in task alignment. Classroom assessments are closely linked to ESP content and communicative tasks practised during the course, whereas the Cambridge test evaluates more general language proficiency under standardised conditions. As a result, classroom performance may better reflect context-specific communicative competence, while standardised scores capture broader, more formal aspects of language ability. In addition, the sensitivity of each measure to short-term instructional effects may differ, which could account for the uneven patterns observed.

As a result, improvements in task-specific fluency and terminology use may occur without movement between CEFR levels, which typically reflect broader, cumulative development. The absence of substantial CEFR shifts therefore does not contradict classroom evidence of progress; rather, it suggests strengthening of domain-specific communicative performance within existing proficiency bands.

Notably, many A1 students achieved good or satisfactory end-of-semester results, while B1 students more frequently attained very good grades. This pattern indicates that classroom achievement reflected professional task competence rather than strict alignment with global CEFR levels.

Digital Readiness

The questionnaire also asked students to reflect on their digital literacy and experience with technology through a set of targeted items addressing

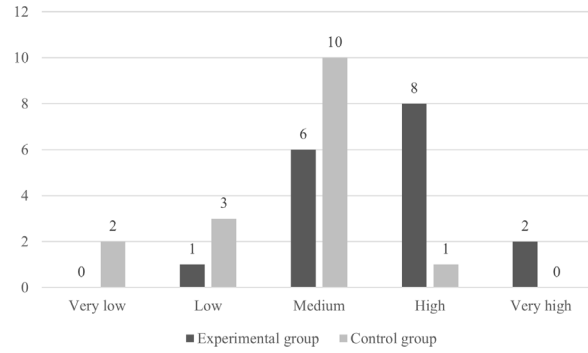
self-assessed computer proficiency, familiarity with digital learning environments, prior exposure to AI-based tools and use of specific educational platforms.

The responses indicate a moderate to high level of perceived computer competence, although the distribution differed across groups (see Figure 7). In the EG, ratings clustered primarily at the high level, with eight students selecting this option and a smaller number indicating very high competence (two students). Fewer participants reported medium or low proficiency and none assessed their skills as very low. By contrast, the CG showed a stronger concentration at the medium level (ten students), alongside a smaller number reporting high proficiency and a limited presence of low or very low ratings.

The above patterns suggest that most respondents felt generally comfortable with basic computer use, while also pointing to meaningful variation in perceived digital competence between groups. This variability provides important context for interpreting students' engagement with AI-supported learning tools later in the study.

Figure 7

Self-Rated Computer Skills by Group (EG vs CG)



Source: Author's research.

Experience with AI Tools

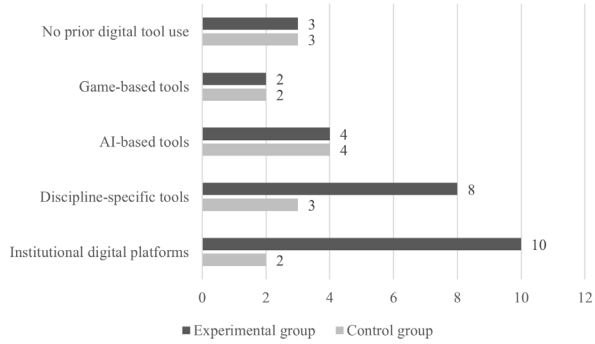
Students reported using a range of digital tools for learning, most commonly institutional platforms (e.g., LMS systems), followed by discipline-specific applications and simulation tools. Game-based and language-learning applications were mentioned less frequently and appeared to function mainly as supplementary resources.

A subset of respondents indicated prior experience with AI-based tools, while others reported limited or no exposure. Overall, participants indicated moderate to high digital literacy, though familiarity with AI-driven systems varied

across the cohort (see Figure 8). This variation provides context for interpreting subsequent findings on readiness and expectations.

Figure 8

Prior Experience with AI-Based Tools (EG vs CG)



Source: Author's research.

Perceptions of AI in Language Learning

This section examines changes in students' attitudes toward AI-supported language learning between baseline and endline (Figure 9). Only the experimental group (EG) engaged with the avatar through structured, teacher-mediated activities, whereas the control group (CG) followed the standard syllabus without guided exposure.

At baseline, both groups expressed generally positive but qualified views. AI was viewed most favourably as a tool for learning medical terminology, while its role in developing communicative skills generated more mixed responses. The idea of AI replacing human instruction was clearly rejected from the outset.

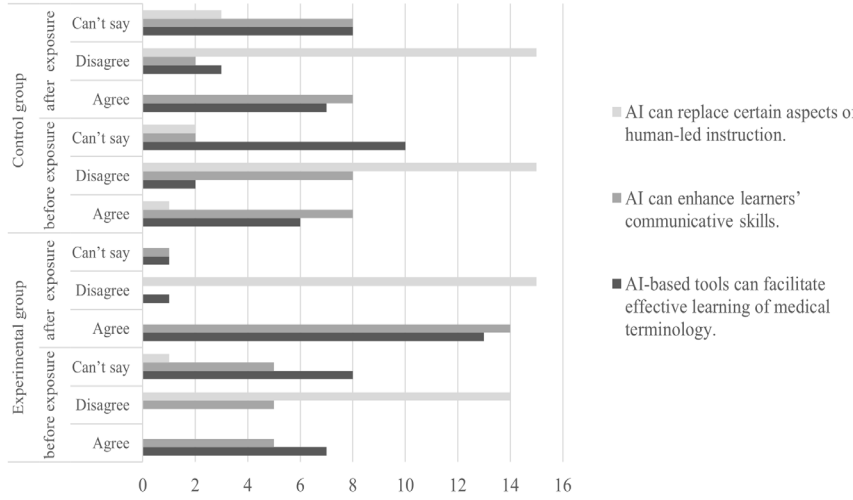
Following sustained classroom integration, the EG showed a marked increase in agreement regarding AI's support for both terminology and communicative development, accompanied by a sharp decline in uncertainty. At the same time, resistance to AI replacing teachers remained strong. This pattern suggests a more differentiated understanding of AI as a supportive rather than autonomous instructional tool.

By contrast, the CG exhibited only modest attitudinal shifts. Although overall views remained broadly positive, uncertainty persisted and confidence in AI's communicative benefits did not increase substantially. Across both groups, students consistently endorsed AI as a supplementary resource while rejecting its role as a substitute for human instruction.

Overall, structured engagement with the avatar appears to have strengthened confidence in its practical value for ESP communication, without diminishing the perceived centrality of the teacher.

Figure 9

Changes in Perceptions of AI in Language Learning (EG vs CG, Baseline and Post-Intervention)



Source: Author's research.

Functionality Expectations for the AI Avatar

Students evaluated the relevance of four avatar functionalities related to professional language use: clinical communication simulation, terminology support, error correction with feedback and pronunciation practice (see Figure 10). Responses were collected at baseline and endline, allowing shifts in both agreement and certainty to be observed.

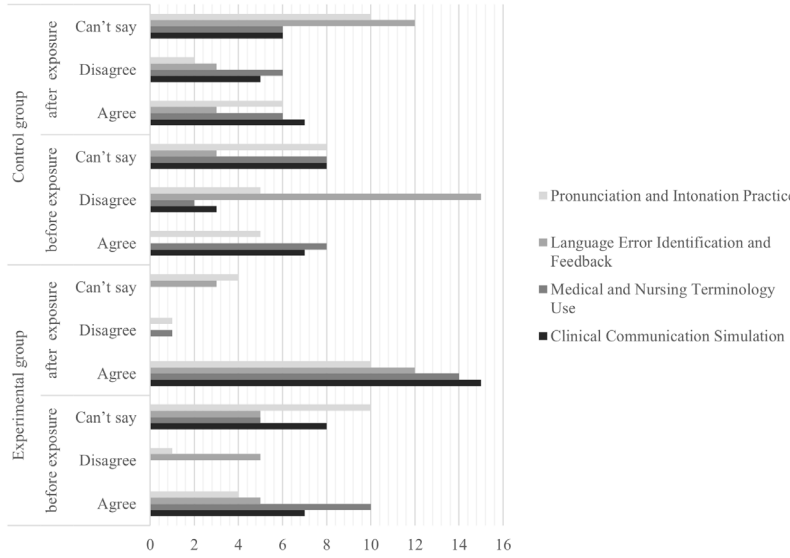
At baseline, the EG already viewed clinical simulation and terminology use favourably, though uncertainty was evident, particularly regarding pronunciation and feedback functions. Following sustained classroom integration, agreement increased across all domains and uncertainty declined markedly. Clinical simulation emerged as the most consistently endorsed feature, while supporting functions gained clearer recognition. Disagreement remained minimal, indicating more experience-based judgements.

The CG displayed a more stable pattern. Although agreement rose slightly over time, uncertainty persisted and no clear consolidation of priorities emerged.

Across both groups, clinical simulation was rated most highly. Overall, the results suggest that guided engagement contributed to a refinement—rather than simple amplification—of expectations regarding avatar functionality (Figure 10).

Figure 10

Perceived Importance of Avatar Functionalities (EG vs CG, Baseline and Post-Intervention)



Source: Author's research.

Design Expectations for the AI Avatar

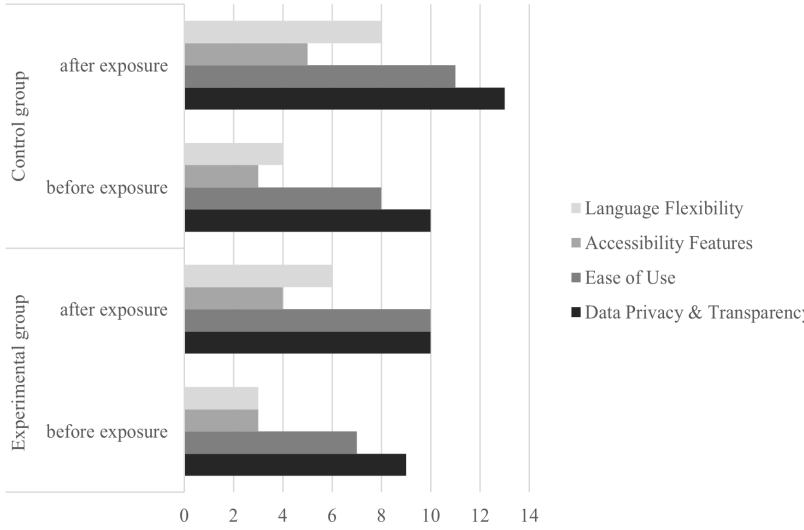
Responses to design-related items indicate that students' expectations became more experience-informed over time (see Figure 11). Across both groups, data privacy and transparency were consistently rated as the most important features, with endorsement increasing after the instructional period—particularly in the CG—suggesting that trust remained a central concern.

Ease of use gained prominence following exposure, especially in the EG, where interface clarity and operational simplicity emerged as explicit conditions for sustained engagement. Accessibility features, although less strongly endorsed overall, showed a modest increase in importance across both groups. A similar pattern was observed for language flexibility (e.g., Polish–English interaction), which became more visible after exposure, particularly in the CG.

Overall, design expectations evolved in a differentiated manner rather than along a simple upward trajectory. Transparency and usability remained central, while accessibility and language flexibility gained recognition without displacing core priorities (Figure 11).

Figure 11

Changes in Design-Related Expectations for the AI Avatar (EG vs CG, Baseline and Post-Intervention)



Open-Ended Questions

Drawing on additional questionnaire items (see Table 1), students were asked about their willingness to participate in future AI-avatar pilots, preferred frequency of use, desired features and implementation concerns.

Willingness increased over time. At baseline, around three in five students expressed readiness to join a future pilot, with many unsure. By the end, approximately four in five were willing, uncertainty declined and refusals were rare. Adoption, however, remained conditional, particularly on clear data-protection and transparency measures.

Preferred frequency also shifted. Initial preference for irregular use gave way to more structured engagement: weekly practice became more common and a small subgroup reported daily use. This suggests that familiarity and perceived gains in terminology and clinical communication supported more regular integration.

Feature expectations evolved from core functions (clinical simulations, terminology support, feedback, pronunciation, bilingual interaction, transcript access) toward more advanced capabilities, including adaptive branching scenarios, richer analytic feedback, enhanced bilingual flexibility, accessibility options, LMS integration and user-controlled privacy settings.

Concerns became more specific over time. Early general reservations about privacy, reliability, usability, teacher replacement and time constraints developed into operational expectations regarding data storage, consent procedures, content accuracy safeguards, bias control and continued human oversight. While usability concerns diminished with experience, time pressure remained a practical constraint.

By and large, responses indicate a shift from cautious curiosity to higher, though conditional, acceptance, accompanied by growing expectations for transparency, pedagogical integration and alignment with authentic clinical communication needs (Table 1).

Table 1

Predicted Shifts from Start to End of the Study (N = 33/38)

Area	Start of the research (33)	End of the research (38)	Interpretation of shift
Willingness to participate in future pilots	Yes: 60.6% Unsure: 33.3% No: 6.1%	Yes: 81.6% Unsure: 13.1% No: 5.3%	↑ Greater readiness after hands-on use; conditional hesitations remain but shrink if privacy/ transparency are addressed
Preferred frequency of use	Irregular: 81.8% Weekly: 12.1% Daily: 6.1%	Irregular: 55.3% Weekly: 36.8% Daily: 7.9%	→ Shift from ad-hoc to more scheduled practice as perceived gains accrue; small motivated subgroup moves to daily use
Desired features (open-ended)	Clinical simulations; medical/nursing terminology; error correction & grammar; pronunciation/intonation; PL↔EN translation (ask in Polish/answer in English); saveable transcripts; clear, well-paced voice; (for some) reminders/group sessions	More advanced & integrated: branching/adaptive clinical scenarios with cultural sensitivity & patient-emotion cues; richer feedback (explanations, examples, progress analytics); stronger bilingual flexibility (code-switching, captions/subtitles) & accessibility (font, captions, TTS); LMS integration, exportable logs, optional offline mode, user-controlled privacy toggles	↑ Expectations evolve from core practice tools to adaptive, data-rich and integrated ecosystems supporting professional communication

Table 1 Continued

Area	Start of the research (33)	End of the research (38)	Interpretation of shift
Concerns & reflections about implementation	Data privacy/transparency; reliability/technical stability; usability/self-efficacy; ethics & fear of replacing human teachers; time constraints	More specific & operational: privacy (storage location, retention/deletion, consent granularity, audit trails); quality (accuracy thresholds for clinical content, bias/cultural safety, guardrails against hallucinations); practice (human oversight, clear policies); usability fears ease with onboarding; time pressure persists	

Discussion

This study examined whether integrating AI-generated avatars into a coursebook-based ESP syllabus influences nursing students' perceptions, preferences and readiness to adopt avatar-supported learning and how these relate to professional English communication. The findings indicate that structured, teacher-mediated integration was associated with stronger attitudinal consolidation and greater perceived preparedness for clinical communication. These results are consistent with research suggesting that AI tools, when pedagogically framed, can reduce language anxiety, promote learner autonomy and provide meaningful rehearsal opportunities for complex interactions (Alsager, 2024; Rezai, Soyoof & Reynolds, 2024).

A central finding concerns the dynamic nature of readiness. At baseline, students expressed cautious openness: AI was viewed as potentially useful, particularly for terminology learning, but not as a substitute for human instruction. Following sustained, guided interaction, the experimental group (EG) revealed clearer endorsement of the avatar's role in supporting communicative development, alongside a continued rejection of teacher replacement. This pattern aligns with Zhang et al. (2025), who link AI literacy and repeated exposure to increased willingness to communicate and reduced anxiety. Readiness, therefore, appears not as a fixed disposition but as an experience-sensitive construct shaped by scaffolding and contextualisation.

Improvements were most visible at the level of domain-specific performance rather than global CEFR advancement. The absence of substantial movement across CEFR bands does not negate classroom gains. Instead, it underscores the distinction between general proficiency and situated professional competence.

Many A1 students achieved satisfactory or good results within the ESP syllabus, suggesting that task-specific communicative development can occur within stable CEFR profiles. This distinction is crucial in professionally oriented language education, where readiness for clinical interaction does not necessarily coincide with measurable shifts in overall proficiency.

The contrast between the EG and the control group (CG) further clarifies the role of pedagogy. In the EG, the avatar was embedded into structured classroom sequences, aligned with coursebook objectives and reinforced through teacher-led discussion. Over time, this integration strengthened students' confidence in the avatar's practical value and increased willingness to use it regularly. In the CG, where access was self-directed and unmediated, attitudinal shifts were modest and uncertainty persisted. Vocabulary and comprehension improved through regular coursework, yet opportunities for scenario-based rehearsal were limited. These findings suggest that technology alone does not reshape perceptions or readiness; its educational impact depends on instructional design and guided use (Zhu & Wang, 2025; Zhang et al., 2025).

This does not imply a diminished role for the instructor. On the contrary, the findings reinforce arguments that AI in education should remain human-centred (Pegrum, 2026; Chiu et al., 2023). Teacher mediation appeared to amplify, rather than replace, pedagogical authority by framing the avatar as a complementary tool. Students consistently endorsed supportive functions—terminology practice, rehearsal, feedback—while rejecting AI as an autonomous instructional agent. Their responses reflect a differentiated appraisal of AI's value, rather than uncritical acceptance.

The concerns regarding data privacy, transparency and reliability remained prominent and became more operational over time. Students increasingly articulated expectations concerning data storage, consent procedures and safeguards against inaccurate or biased content. Such concerns are especially pertinent in healthcare contexts, where trust and ethical responsibility are central (Chiu et al., 2023; Peña-Acuña & Durão, 2024). Addressing these issues requires institutional policy frameworks and explicit communication about data governance. As Liu (2025) argues, sustainable AI integration depends not only on technical functionality but also on ethical literacy among educators and learners.

Several implications follow. First, avatar-supported learning appears most effective when integrated into syllabus-based instruction and aligned with clearly defined professional communication outcomes. Second, iterative rehearsal combined with feedback may strengthen confidence and functional fluency within existing proficiency bands. Third, institutional adoption should prioritise transparency, usability and human oversight to maintain trust and pedagogical coherence.

In sum, the pedagogical contribution of AI avatars in nursing ESP lies not in transforming global proficiency levels within a single semester, nor in replacing

instructors, but in extending opportunities for structured, scenario-based practice. When embedded within human-guided instruction and supported by ethical safeguards, such tools may enhance students' readiness for authentic clinical communication while preserving the central role of the teacher.

Conclusions

This study examined how the integration of an AI-generated avatar within a nursing ESP course relates to students' perceptions, readiness and communicative development. The findings suggest that structured, teacher-mediated use is associated with more consistent attitudinal alignment and clearer indications of domain-specific communicative development than self-directed exposure alone.

The results point to a distinction between general language proficiency and task-specific competence. While no substantial movement across CEFR levels was observed, classroom evidence indicates strengthening of functional communication within professional contexts. This suggests that avatar-supported learning may be particularly relevant for reinforcing situated communicative performance rather than accelerating global proficiency gains.

A central implication concerns instructional design. The effectiveness of the avatar appears contingent on its integration within guided classroom practice, where interaction is scaffolded and linked to professional tasks. In this sense, the tool functions not as an autonomous instructional agent but as a structured environment for rehearsal and feedback.

At the same time, the findings remain exploratory. The sample is limited, the design is quasi-experimental and the analysis is primarily descriptive. Further research using larger samples and inferential methods would be required to determine the extent to which these patterns can be generalised.

References

Alsager, H. N. (2024). To look from another window in education: Artificial intelligence assisted language learning and its reflections on academic demotivation, foreign language learning anxiety and autonomy. *Computer-Assisted Language Learning Electronic Journal, 25*(4), 124-147. https://callej.org/index.php/journal/article/view/469

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge University Press.

Chen, C.-H., Koong, C.-S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support students' English speaking skills, learning anxiety and cognitive load. *Educational Technology & Society, 25*(1), 1–14. https://doi.org/10.30191/ETS.202201_25(1).0001

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 4*, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford University Press.

Evans, V., Dooley, J., & Wright, R. (2019). *Career paths: Nursing – Student's book*. Express Publishing.

Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge.

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

Liu, W. (2025). Language teacher AI literacy: Insights from collaborations with ChatGPT. *Journal of China Computer-Assisted Language Learning, 5*(2), 287–316. https://doi.org/10.1515/jccall-2024-0030

Pegrum, M. (2026). From revolution to evolution: What generative AI really means for language learning. *Language Teaching, 59*(2), *180–196.* https://doi.org/10.1017/S0261444825000151

Peña-Acuña, B., & Corga Fernandes Durão, R. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education*, *9*, Article 1490067. https://doi.org/10.3389/feduc.2024.1490067

Rezai, A., Soyoof, A., & Reynolds, B. L. (2024). Disclosing the correlation between using ChatGPT and well-being in EFL learners: Considering the mediating role of emotion regulation. *European Journal of Education, 59*(4), Article e12752. https://doi.org/10.1111/ejed.12752

Roe, J., Perkins, M., & Furze, L. (2025). From assessment to practice: Implementing the AIAS framework in EFL teaching and learning [Preprint]. British University Vietnam.

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3–20). Logos International.

Woo, D. J., Yu, Y., & Guo, K. (2025). Exploring EFL secondary students' AI-generated text editing while composition writing. *arXiv.* https://doi.org/10.48550/arXiv.2505.17041

Zhang, Q., Nie, H., Fan, J., & Liu, H. (2025). Exploring the dynamics of artificial intelligence literacy on English as a Foreign Language learners' willingness to communicate: The critical mediating roles of artificial intelligence learning self-efficacy and classroom anxiety. *Behavioral Sciences*, *15*(4), Article 523. https://doi.org/10.3390/bs15040523

Zhu, M., & Wang, C. (2025). A systematic review of AI in language education: Current status and future implications. *Language Learning & Technology*, *29*(1), 1–29. https://doi.org/10.64152/10125/73606

Appendix 1

Artificial Intelligence in English Language Learning

This questionnaire explores respondents' views on the use of an AI-generated avatar (developed with the Soul Machines platform) to support English language learning, including the development of communicative skills, acquisition of medical terminology and flexible learning independent of time and location. Participation is anonymous and all data will be reported only in aggregated form. There are no right or wrong answers and all responses are valuable for the study. Please read each question carefully and answer honestly. The questionnaire takes approximately 10–15 minutes to complete. No personal data, such as name or email address, will be collected unless you choose to provide it voluntarily.

General information

1. Age
   18–21
   22–25
   26 and above
2. Gender
   Female
   Male
   Prefer not to answer
3. Level of Study
   First-cycle studies (Bachelor's/Engineering degree)
   Second-cycle studies (Master's degree)

Digital Tools

4. How would you assess your overall level of computer literacy?
   Very high
   High
   Moderate
   Low
   Very low

5. How would you assess your level of familiarity with digital tools used for learning?
   Very low
   Low
   Moderate
   High
   Very high
6. Which digital tools or platforms have you used or are you currently using for learning purposes? (e.g., Moodle, Khan Academy, simulation laboratories, educational platforms, medical applications)
7. Have you previously used AI-based (artificial intelligence) tools, such as ChatGPT, Duolingo, or medical applications?
   Yes
   No
   Yes, but I do not recall which ones

Level of English Language Proficiency (Based on the Completed Test)

8. What is your level of English language proficiency?
   A1
   A2
   B1
   B2
   C1
   C2

AI in learning English

9. To what extent do you agree with the following statements?

	Strongly agree	I agree	Hard to say	Disagree	Strongly disagree
AI-based tools can facilitate effective learning of medical terminology.					
AI can enhance learners' communicative skills.					
AI can replace certain aspects of human-led instruction.					

10. What are your initial thoughts on the use of AI in English language learning?
    (Open-ended question)

11. What concerns, if any, do you have regarding the use of AI in English language learning?
(e.g., reliability, ethical issues, data privacy)
(Open-ended question)
12. In which tasks should an AI avatar provide support? (Multiple responses may be selected)
Learning medical terminology
Simulating patient–provider communication scenarios
Providing reminders about tasks or deadlines
Answering questions related to learning or professional tasks
Offering career-related or professional guidance
13. How important is it that the AI avatar:
For each of the following statements, please indicate your response using the scale below:
1. definitely not important 2. rather not important 3. difficult to say
4. rather important 5. definitely important

	1	2	3	4	5
Uses accurate medical and nursing terminology.					
Simulates clinical scenarios to practice communication (**e.g., interactions with patients or colleagues**).					
Identifies and corrects language errors and suggests appropriate forms.					
Provides exercises focused on pronunciation and intonation.					
Incorporates cultural sensitivity in patient-related scenarios.					

14. Do you believe that an AI avatar could help you prepare for clinical placements conducted in English?
Yes
No
Not sure
15. Which aspects of English language learning would you like to develop with the support of an AI avatar?
(Open-ended question)

Ethics and privacy

16. Would you feel comfortable using an AI avatar if your data were collected for the purpose of improving its performance?
Yes
No
Depends on transparency in the use of data

17. How important is it that the AI avatar explains how your data are used?
    Not important
    Of little importance
    Difficult to say
    Important
    Very important

Availability and technical requirements

18. Do you have access to the devices and internet connectivity required to use an AI avatar?
    Yes
    No
19. Do you require any accessibility features? *(e.g., text-to-speech, larger font size)*
    Yes
    No

Willingness to participate in the pilot study

20. Would you be willing to participate in English language learning using an AI avatar in the future?
    Yes
    No
    Hard to say
21. How frequently would you like to use an AI avatar during your learning process?
    Daily
    Weekly
    Whenever time permits

Opinions and Suggestions

22. What features or capabilities would you like the AI avatar to have in order to support your English language learning?
    (Open-ended question)
23. Do you have any concerns or suggestions regarding the implementation of an AI avatar in English language learning?
    (Open-ended question)

Appendix 2

Artificial Intelligence in English Language Learning

This questionnaire explores respondents' views on the use of an AI-generated avatar (developed with the Soul Machines platform) to support English language learning, including the development of communicative skills, acquisition of medical terminology and flexible learning independent of time and location. Participation is anonymous and all data will be reported only in aggregated form. There are no right or wrong answers and all responses are valuable for the study. Please read each question carefully and answer honestly. The questionnaire takes approximately 10–15 minutes to complete. No personal data, such as name or email address, will be collected unless you choose to provide it voluntarily.

1. To what extent do you agree with the following statements?

	Strongly agree	I agree	Hard to say	Disagree	Strongly disagree
AI-based tools can facilitate effective learning of medical terminology.					
AI can enhance learners' communicative skills.					
AI can replace certain aspects of human-led instruction.					

2. What are your initial thoughts on the use of AI in English language learning? *(Open-ended question)*
3. What concerns, if any, do you have regarding the use of AI in English language learning?
   (e.g., reliability, ethical issues, data privacy)
   (Open-ended question)
4. In which tasks should an AI avatar provide support? *(Multiple responses may be selected)*
   Learning medical terminology
   Simulating patient–provider communication scenarios
   Providing reminders about tasks or deadlines
   Answering questions related to learning or professional tasks
   Offering career-related or professional guidance

5. How important is it that the AI avatar:
   For each of the following statements, please indicate your response using the scale below:
   1. definitely not important 2. rather not important 3. difficult to say 4. rather important 5. definitely important

	1	2	3	4	5
Uses accurate medical and nursing terminology.					
Simulates clinical scenarios to practice communication (**e.g., interactions with patients or colleagues**).					
Identifies and corrects language errors and suggests appropriate forms.					
Provides exercises focused on pronunciation and intonation.					
Incorporates cultural sensitivity in patient-related scenarios.					

6. Do you believe that an AI avatar could help you prepare for clinical placements conducted in English?
   Yes
   No
   Not sure
7. Which aspects of English language learning would you like to develop with the support of an AI avatar?
   (Open-ended question)

Ethics and privacy

8. Would you feel comfortable using an AI avatar if your data were collected for the purpose of improving its performance?
   Yes
   No
   Depends on transparency in the use of data
9. How important is it that the AI avatar explains how your data are used?
   Not important
   Of little importance
   Difficult to say
   Important
   Very important

Availability and technical requirements

10. Do you have access to the devices and internet connectivity required to use an AI avatar?
    Yes
    No

11. Do you require any accessibility features? (*e.g., text-to-speech, larger font size)*
    Yes
    No

Willingness to participate in the pilot study

12. Would you be willing to participate in English language learning using an AI avatar in the future?
    Yes
    No
    Hard to say
13. How frequently would you like to use an AI avatar during your learning process?
    Daily
    Weekly
    Whenever time permits

Opinions and Suggestions

14. What features or capabilities would you like the AI avatar to have in order to support your English language learning?
    (Open-ended question)
15. Do you have any concerns or suggestions regarding the implementation of an AI avatar in English language learning?
    (Open-ended question)