Journal of Curriculum Studies Research

https://curriculumstudies.org

E-ISSN: 2690-2788

Volume: 7 Issue: 2 2025

pp. 31-52

AI as a Simulated Client: Impact on Educational Psychology Students' Therapeutic Skills

Dinara Kakabayeva[a], Aizhana Abibulayeva[b], Kuldarkhan Orazbayeva[a] & Liza Naviy[a]

* Corresponding author
Email: lnaviy@shokan.edu.kz
a. Department of Pedagogy and Psychology, Ualikhanov University, Kokshetau, Kazakhstan
b. Department of Pedagogy, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

Article Info
Received: January 21, 2025
Accepted: June 22, 2025
Published: August 10, 2025

10.46303/jcsr.2025.10

How to cite
Kakabayeva, D., Abibulayeva, A., Orazbayeva, K. & Naviy, L. (2025). AI as a Simulated Client: Impact on Educational Psychology Students' Therapeutic Skills. *Journal of Curriculum Studies Research*, 7(2), 31-52. https://doi.org/10.46303/jcsr.2025.10

ABSTRACT
Role-playing emerges as a central pedagogical strategy, enabling trainees to navigate complex scenarios through human-like interactions that bridge theoretical knowledge with practical expertise. As generative artificial intelligence continues its exponential growth, its promise in mental health contexts has garnered substantial attention. However, the integration of chatbots into educational psychology training remains an underexplored frontier. This study aims to address this gap by examining the impacts of an intervention using role-play simulations of psychoeducational counseling sessions on therapeutic skills. This pre-posttest controlled study engaged 112 educational counseling students. A treatment group (n = 53) participated in an eight-week intervention using a Gemini-powered simulation, which featured a pre-prompted student-profile chatbot and an artificial third party providing post-activity feedback on participants' performance as counselors, while a comparison group (n = 59) had no access to the simulation. Therapeutic skill assessments were conducted before and after the experiment, complemented by qualitative interviews exploring students' perceptions. The chatbot-mediated intervention yielded gains in feedback synthesis, guided discovery, cognitive-behavioral focus, and change strategies. While some respondents expressed skepticism, overall sentiment was positive. This research highlights the potential of chatbot-augmented training to enhance empathetic understanding, interrogative precision, and reflective feedback capabilities in educational psychology curricula.

KEYWORDS
Counseling; Gemini; generative artificial intelligence; LLM; role-play.

INTRODUCTION

The incorporation of recently introduced web-based solutions into societal systems and educational paradigms (Akhmetov et al., 2025; Sudirman et al., 2025; Turarbek et al., 2023) has redefined the boundaries of experiential learning, particularly in disciplines requiring nuanced interpersonal skills such as counseling. Innovative instructional approaches are now being pursued with renewed vigor, aiming to address the enduring challenge of balancing theory with hands-on practice (Koishybekova et al., 2025; Sun et al., 2025). Mental health provider education programs commonly rely on a synthesis of theoretical instruction and practical engagement (Green & Adawi, 2024), with the latter often delivered through role-play simulations with peers. Role-playing, which immerses trainees in specific roles within realistic simulated scenarios mirroring human interactions, is widely recognized as a method for practicing in front of others and refining problem-solving skills across domains (Chen et al., 2025; Darmawansah et al., 2025). This method offers learners the opportunity to navigate intricate cases in a controlled yet dynamic setting, thereby consolidating theoretical understanding with practical skill (Terzi et al., 2025). In counselor education, it serves as a cornerstone for honing both micro counseling skills, such as active listening and perception checking, and advanced therapeutic techniques (Coohey et al., 2025; Grunhaus & Martin, 2025; Horton & Lloyd-Hazlett, 2025). These micro skills, essential for establishing a strong therapeutic alliance, are widely accepted as the foundation of effective counseling (Maurya, 2024a; Muryono, 2024). Triad exercises, in which students take on the roles of therapist, client, or observer, exemplify the role-play approach (Maurya, 2024b). This framework enables rotated skill practice and the receipt of targeted feedback. However, such methods are resource-intensive, spatially limited, and dependent on the varying proficiency of human participants, often leading to inconsistent learning experiences (Barker et al., 2024; Chen et al., 2025; Shaleh et al., 2022).

In light of the recent advances in computer science, the academic community has increasingly focused on the transformative potential of artificial intelligence (AI) in learning and teaching contexts (Adipat, 2025; Tan et al., 2025), as well as in specific sectors like AI-augmented art therapies (Luo et al., 2024). However, this enthusiasm is balanced by ongoing debates about the ethical implications and potential risks of AI integration in educational settings (El Khayati et al., 2025). Concerns have emerged regarding algorithmic bias (Nguyen, 2025), data privacy (Symeou et al., 2025), potential over-reliance on AI systems (Alshamy et al., 2025), and the risk of diminishing human connection in fields that rely heavily on interpersonal relationships (Alotaibi, 2025). It can thus be argued that AI-mediated training may unintentionally promote a mechanistic approach to counseling, potentially weakening the development of authentic empathy and intuitive therapeutic judgment. Additionally, questions remain about the accountability and transparency of AI-generated feedback, especially when training future mental health professionals who will work with vulnerable populations. Large language models (LLMs), such as chatbots like ChatGPT and Gemini, have shown considerable promise as dialogue partners capable of generating responsive learning environments (Baytak, 2024; Ibrahim &

Ajlouni, 2024; Murtaza et al., 2025; Zheng et al., 2025). A key strength of these generative AI (GenAI) systems, built on advanced natural language processing, is their ability to deliver case-specific, automated feedback, a critical element of effective learning (Daniel et al., 2025; Mi et al., 2025; Nofal et al., 2025). While earlier AI-based feedback systems required extensive resources, including labor-intensive annotation for task-specific training datasets (Karjus, 2025), modern self-supervised conversational agents overcome such barriers. By using prompt-based adaptability without the need for custom tuning, these models generate nuanced, contextually relevant feedback, as validated by end-users and domain experts (Kinder et al., 2025).

Employing a generative assistant as a virtual client for practicing and assessing counseling skills may reduce the discomfort associated with human role-playing, such as performance anxiety caused by observation from peers or instructors (Baker & Jenney, 2023). LLM chatbots are designed to emulate a wide range of interlocutors, approaching the complexity of human interactions (Yusuf et al., 2025). Despite these advantages, certain limitations of GenAI agents remain. Their inability to detect and convey non-verbal cues, along with their reliance on pre-programmed response patterns, may fail to fully capture subtle emotional nuances (Chan, 2025).

LITERATURE REVIEW

The rapid rise of general-purpose LLMs has driven advancements across various sectors, including education and healthcare, due to their broad accessibility and advanced functionalities. Specifically, the potential use of conversational AI in mental health contexts has been documented in numerous studies, highlighting its promise in enhancing diagnostic precision, emotion recognition, and risk assessment (Haber et al., 2025), indicating its potential to reshape therapeutic practices.

Despite this growing interest, the integration of generative chatbots into psychological counseling training within educational settings remains largely unexplored, both in applied and theoretical domains. This gap is especially important given the critical need for experiential learning in counselor education, where repeated practice and immediate feedback are essential for skill mastery. The scholarly literature so far has been dominated by theoretical discussions (Krach & Corcoran, 2024; Nelson et al., 2025), reviews (Chen et al., 2024a), and studies examining stakeholder perceptions (Arnout & Alshehri, 2025; Kuhail et al., 2024). Furthermore, the focus of these publications often centers on indirect indicators of GenAI's value in mental health service provision. For example, a systematic review (Levkovich & Omar, 2024) of 29 studies from 2020 to 2024 examined the use of LLMs in suicide prevention by identifying depression and suicidal ideation from various sources, such as electronic health records and social media posts. The review found that some models demonstrated impressive encoding accuracy, often matching or surpassing mental health professionals in suicide risk assessment.

When the analytical focus is narrowed to psychologist-client interactions, the lack of research becomes even more evident, with most efforts emphasizing the use of generative AI as a therapist rather than as a client. For example, a mixed-methods study (Vowels et al., 2024)

involved 20 participants who received single-session relationship advice generated by ChatGPT. Participants gave high ratings to the chatbot's attributes, including its perceived human-likeness and its clarity and actionable steps for addressing relationship issues. In contrast, another recent study (Eryilmaz & Başal, 2024) compared the emotional reflections of human counselors and ChatGPT in response to client statements within therapeutic dialogues. Both quantitative and qualitative data led the authors to argue that the AI model cannot replace the unique advantages of real psychologists, particularly regarding personal resonance and emotional depth.

Nevertheless, preliminary findings from studies using generative chatbots as simulated clients are promising. In a proof-of-concept trial (Fung & Laing, 2024), one of the authors acted as a cognitive behavioral therapist, while ChatGPT 4 was prompt-engineered to represent a client with depression. The interactive agent appropriately responded to all the researcher's inquiries and techniques. Similarly, in another study (Maurya, 2024a), the author took on the role of counselor, while ChatGPT portrayed a client across 10 pre-designed vignettes. The results indicated that the dialogue system exhibited consistency and appropriate emotional expression in its client portrayal. One limitation noted was that LLM responses were sometimes overly formulaic, lacking the spontaneous variability typically seen in real-world individual behavior.

To the best of our knowledge, only one study to date reports on the structured, longitudinal use of GenAI as a simulated client for mental health provider training. The study (Prescott et al., 2024) employed chatbot technology to create an emotionally responsive interactive client designed to introduce student therapists to the core principles of person-centered therapy through a series of pre-programmed scenarios. Participants noted the realism and engagement of the scenarios, with many describing the experience as similar to conversing with a real person. However, the study was limited by its reliance on qualitative data.

Given the demonstrated yet underexplored potential of chatbots to address this gap, rigorous empirical research in this area is both timely and necessary. Building on the LLM-based role-play triad framework proposed by Prescott et al. (2024), the present study implements a custom-instructed Gemini chatbot to serve both as a virtual client and as a source of automated corrective feedback. By structuring a triadic interaction between a human student, AI-simulated client, and AI-embodied reviewer, the intervention aims to measure improvements in therapeutic skills from pre-test to post-test and to capture participants' perceptions of this training approach.

This study seeks to empirically assess the effectiveness of a Gemini-powered role-playing intervention in enhancing therapeutic skills among educational counseling students. The research is guided by the following questions:

- Does participation in the chatbot-mediated simulation lead to significant improvements in therapeutic skills compared to a business-as-usual group?
- How do students perceive the chatbot-augmented training experience?

This study addresses a critical gap in the literature, providing evidence on the effectiveness and acceptability of chatbot-augmented training in educational psychology, a topic of growing importance as AI technologies become more integrated into professional education. Empirically examining the efficacy of GenAI-facilitated role-playing offers insights that can guide future applications of LLM-assisted interventions in mental health education.

MATERIALS AND METHODS

Design and Sample

This study used a quasi-experimental, pre-posttest mixed design with a negative control group, a model commonly applied in educational and psychological intervention research to assess the causal impact of an intervention when random assignment is not possible (Hamed et al., 2025). The research began in early September 2024 with participant recruitment after obtaining informed consent and project approval from the institutional ethics committee, and concluded in late February 2025 with data analysis and interpretation. The sample included 112 undergraduates pursuing educational counselor diplomas in their third and fourth years at the corresponding author's university. Participants had an average age of 20.6 years and were predominantly female (94 out of 112 individuals).

Intervention Procedures

This eight-week course immersed educational counseling students in an interactive experiment. In the experimental group (n = 53), students engaged one-on-one with a conversational chatbot designed to simulate a counseling client, creating a dynamic training ground within a counseling triad. In this setup, the student acted as a counselor, while the chatbot served both as a virtual client presenting issues and as an observant tutor providing post-session critiques on counseling performance. As a core element of the experiment, these students participated in weekly text-based role-playing sessions with the artificial persona at dedicated computer stations on campus after their formal lectures. Meanwhile, the control group (n = 59) followed a standard curriculum and did not engage in virtual practice sessions or receive AI-generated feedback, serving as a baseline for comparison. Prior to the entry assessment, the sample was balanced at 64 participants per group using randomization software. However, uneven group sizes resulted when some participants did not complete all required procedures.

Participants accessed Gemini 2.0 using credentials provided by research assistants, a measure taken to protect their personal data from Google. They were explicitly instructed not to share any potentially identifying information with the chatbot. The protocol required participants to type their responses, initiating each conversation with a simple greeting. This action triggered the custom LLM profile to randomly present one of eight carefully adapted case scenarios from Maurya (2024a) (see Appendix). The agent was configured to avoid repeating scenarios in future sessions with the same user. Trainees were tasked with reflecting the simulated client's statements, offering validation and empathy, staying aligned with the client's narrative without over-interpreting, and posing focused follow-up questions to explore the

situation, all while maintaining a natural conversational flow. Ideally, these dialogues lasted about 20 minutes, ending with the student providing relevant strategies for improvement. This advisory stage was intentionally delayed until the ‘‘client’’ had fully explored their concerns and the trainee had a solid understanding of the issue, allowing for personalized suggestions. After the role-play concluded, the GenAI mentor provided feedback on the student’s counseling approach. Human proctors, independent of the research team, monitored the sessions to ensure no communication occurred between participants.

The experiment included eight distinct sessions, aiming for each participant to engage with all eight pre-designed case studies. Prior to full implementation, the role-play mechanics, data collection procedures, and measurement instruments were piloted with 12 undergraduates who were not part of the main study. The intervention began in the first week with an introduction to the learning environment and initial training on corrective feedback, ensuring all participants had a consistent starting point.

Data Collection and Sources

To assess participants’ therapeutic skills at baseline and after the intervention, human-human simulated sessions were conducted with six general psychology graduates acting in various patient roles. The assessments were held over 12 days before and 10 days after the intervention. Each participant completed one session lasting about 10 minutes. The actors were instructed to avoid introducing contradictory details to keep the cases consistent. The simulations were supervised by two licensed therapists and video-recorded by research assistants. After the eight-week role-playing intervention, the videos were independently scored by four post-graduate counselors (blind to the study’s purpose and group assignment) using the skill section of the Inventory of Therapeutic Interventions and Skills (ITIS; Boyle et al., 2020). This observation-based measure assesses 11 micro-skills on a 7-point Likert scale, from 0 (poor) to 6 (excellent). Raters completed Zoom training to standardize their scoring based on the behavioral descriptions in the inventory. Students’ final scores were calculated by averaging the ratings from the four evaluators. Inter-rater reliability for the 11 criteria was substantial (Fleiss Kappa: 0.81 to 0.93).

To ensure transparency and the potential for high scores on the ITIS, the 11 skills were introduced and explained to participants before the baseline measurement. This allowed students to understand what was expected of them for better performance. This approach avoided the flaw noted in the study by Großmann and Krüger (2024), where participants were unable to practice applying the 24 performance criteria between evaluation points due to a lack of awareness. Additionally, the post-interaction AI feedback was aligned with ITIS standards, so the intervention activities were designed to guide pre-service psychologists toward improved outcomes. After the objective post-intervention assessment, impressions of the role-play group’s experience with the training approach were gathered through individual semi-structured, audio-recorded interviews conducted by three journalism graduates. To guide qualitative data collection, four interview questions were adapted from Hinzmann et al. (2023).

Data Analysis

To examine intervention effects while adjusting for potential initial between-group imbalance, ITIS scores were analyzed using 11 separate repeated-measures analyses of covariance (RM ANCOVAs). Pre-test performance for each criterion served as a covariate. Normality and homogeneity of variances were assessed using Shapiro-Wilk and Levene's tests, respectively. If an assumption was violated, an aligned rank transformation analysis of variance was used instead of RM ANCOVA. In both parametric and non-parametric analyses, partial eta squared (η^2p) was used to quantify effect sizes, with thresholds of < 0.06 (small), 0.06–0.13 (medium), and ≥ 0.14 (large). Statistical significance was set at $\alpha < 0.05$. Interview transcriptions were analyzed inductively through thematic analysis using NVivo software. The initial codes, generated automatically, were refined through an iterative process to ensure coding agreement. These refined codes were then grouped into themes.

RESULTS

Quantitative Findings

Prior to inferential analyses, Shapiro-Wilk tests indicated that the assumption of normality was met ($p \geq 0.05$) for the micro-skills of pacing and efficient use of time ($W = 0.965$, $p = 0.055$), use of feedback/summaries ($W = 0.977$, $p = 0.050$), guided discovery ($W = 0.986$, $p = 0.314$), therapeutic relationship/collaboration ($W = 0.975$, $p = 0.064$), and handling problems/questions/objections ($W = 0.941$, $p = 0.059$). However, normality was violated ($p < 0.05$) for clarity of communication ($W = 0.955$, $p = 0.001$), rationale ($W = 0.967$, $p = 0.007$), empathic understanding ($W = 0.957$, $p = 0.001$), focusing on key cognitions and behaviors ($W = 0.960$, $p = 0.002$), strategy for change ($W = 0.959$, $p = 0.002$), and application of techniques ($W = 0.930$, $p = 0.001$). Levene's test results showed that the homogeneity assumption was met ($p \geq 0.05$) for all items except empathic understanding ($F [1, 110] = 7.48$, $p = 0.007$). Descriptive statistics for the therapeutic skills are presented in Table 1 (see appendix), providing an overview of group performance at baseline and after the intervention.

RM ANCOVA showed that for pacing and efficient use of time, no significant group effect was observed ($F[1, 109] = 2.25$, $p = 0.136$), though the effect size was large ($\eta^2p = 0.15$). Similarly, no significant between-group difference was observed for clarity of communication ($F[1, 110] = 0.203$, $p = 0.653$), with a small effect size ($\eta^2p = 0.01$). A significant group effect emerged for use of feedback/summaries ($F[1, 109] = 17.6$, $p < 0.001$), favoring the role-play group, with a large effect size ($\eta^2p = 0.14$), indicating a substantial impact of the intervention. For rationale, RM ANCOVA did not reveal a significant group effect ($F[1, 110] = 0.350$, $p = 0.555$), and the effect size was small ($\eta^2p = 0.01$).

In contrast, a highly significant group difference was found for guided discovery ($F[1, 109] = 40.6$, $p < 0.001$), with post-test scores higher in the treatment group. The effect size was large ($\eta^2p = 0.27$), indicating a strong intervention effect. For therapeutic relationship/collaboration, no significant group effect was detected ($F[1, 109] = 1.45$, $p = 0.230$), with a negligible effect size

(η^2p = 0.01). Similarly, handling problems/questions/objections showed no significant group effect (F[1, 109] = 0.478, p = 0.491), with a small effect size (η^2p = 0.01). For empathic understanding, no substantial group difference was found (F[1, 110] = 3.155, p = 0.785), with a small effect size (η^2p = 0.03). However, a significant between-group difference was observed for focusing on key cognitions and behaviors (F[1, 110] = 4.64, p = 0.033), with better post-test performance in the simulation group. The effect size was small (η^2p = 0.04). Additionally, students in the experimental group showed significantly greater improvement in developing a strategy for change compared to the control group (F[1, 110] = 6.71, p = 0.011), with a medium effect size (η^2p = 0.06). Finally, for application of techniques, the between-group difference approached significance but did not reach the conventional threshold (F[1, 110] = 3.554, p = 0.062), with a small effect size (η^2p = 0.03).

Qualitative Findings

The analysis of participants' post-intervention perceptions of the experience revealed several key areas related to their engagement with the simulated sessions and the perceived value for skill development. Table 2 (see appendix) presents the emergent themes, their definitions, and illustrative quotes from the participants.

Most participants viewed the simulated sessions as a valuable addition to their training. The chance to engage with varied client scenarios in a controlled setting was frequently emphasized. For many, the realism of certain interactions enhanced their engagement. One student reflected on a particularly impactful session: "When 'Juri' talked about the pressure from his family, it felt very real. I could almost picture a student I know saying similar things. It made me really focus on validating his feelings." This sense of realism, when present, appeared to support a more authentic use of counseling skills. However, interacting with an AI also presented challenges. Some students found it difficult at times to maintain a natural conversational flow. One participant shared, "The hardest part was when the responses were a bit off or didn't quite match the nuance of what I was trying to ask. It kind of broke the immersion sometimes, and I had to consciously steer the conversation back on track." These moments highlighted the current limitations of AI in fully replicating human interaction, but they also provided useful opportunities for students to practice redirecting and clarifying in a therapeutic context.

The inclusion of the AI tutor for immediate feedback was viewed very positively. Students appreciated the prompt and specific nature of the feedback. "I loved getting that instant critique," shared one participant. "It was so useful to see exactly where I could have been more empathetic or when my questions were maybe leading." This immediate reinforcement was seen as a strong tool for in-the-moment learning and adjustment of technique. The feedback was not just corrective; it also affirmed positive actions. "It was really motivating when the tutor pointed out that I had done a good job of reflecting or showing empathy. It made me feel like I was getting the hang of it," remarked another student. For many, the simulation offered a much-needed safe space to practice and make mistakes without real-world consequences. "Knowing

it was not a real person meant I felt braver trying out different approaches," one student reflected. "If something did not land well, it was not the end of the world, just a learning opportunity for the next session." This feature was especially helpful for building confidence, particularly when handling more sensitive or complex issues in the scenarios.

The experience also seemed to stay with participants beyond the scheduled sessions. Several students mentioned replaying conversations in their minds or thinking about how they might have handled situations differently. "I kept thinking about the case with Erzhan and his parents' divorce," said one participant. "It made me reflect on the importance of recognizing how family dynamics affect a student's well-being, even if they do not mention it directly." These reflections suggest deeper engagement with the material and ongoing processing of the skills being learned. When asked about difficult moments, students often mentioned either technical limitations of the AI or the emotionally challenging nature of some scenarios. "The session with Sabira was tough," admitted one student. "Her feeling of hopelessness came through, even in text, and it was hard to know the 'right' thing to say without giving false hope." On the other hand, fulfilling moments were often described as times when students felt they had connected with the simulated client or helped the client gain insight. "With Kairat, who was unsure about his future, I felt a real sense of accomplishment when he began to express his own interests more clearly after I asked a few targeted questions. This may sound silly, but it felt like a small victory," shared another participant.

In terms of suggestions for improvement, some students wanted more complexity in the AI's responses and the option for longer, more in-depth conversations. "Maybe in the future, it could allow for more back-and-forth on a single topic, really digging deep instead of moving through things relatively quickly to get to suggestions," suggested one student. Another mentioned wanting "a bit more variation in the way similar emotions were expressed across different clients" to better challenge their ability to tailor responses. Overall, the sentiment toward the AI-driven experience was largely positive, with students recognizing its value in offering practical, low-stakes opportunities to develop and refine key skills. The combination of varied scenarios and immediate, targeted feedback appears to be a promising approach in educational psychology training.

DISCUSSION

This inquiry aimed to determine whether exposing educational counseling students to a chatbot acting as a simulated client and tutor would enhance their therapeutic competencies, as measured in naturalistic settings. The goal was to explore the viability of using a GenAI system to foster counseling micro-skills through repeated role-plays. The data showed that students who participated in the virtual interactive sessions demonstrated stronger post-intervention skills than peers who did not engage in chatbot-mediated exercises. Pre-service psychologists who received chatbot-facilitated practice made notable progress across several skill areas, providing evidence that AI-supported practice can serve as an effective supplementary tool in

counseling education. Higher scores in empathic understanding suggest that repeated structured sessions with the LLM client fostered a greater sensitivity to emotional cues and a more consistent display of warmth toward the client's narrative. Improvement in guided discovery indicates that the repeated practice of crafting open-ended questions in text-based interactions strengthened participants' ability to elicit and explore key cognitions. Growth in the use of feedback/summaries likely stemmed from frequent, immediate chatbot critiques, which taught participants to restate and check comprehension throughout the interaction. In essence, these skill gains align closely with the training platform's goal: to cultivate the ability to validate, probe, and cohesively summarize the simulated client's experiences.

The present quantitative outcomes align with Sharma et al.'s (2023) finding that AI-driven feedback can considerably enhance empathic behaviors during supportive exchanges. While Sharma et al. examined peer-to-peer mental health support and reported nearly a 20% increase in empathic responses, this study observed a similar trend of heightened empathic engagement among educational counselor trainees. Both studies demonstrate that LLM-informed feedback can strengthen interpersonal sensitivity without encouraging over-reliance on algorithmic guidance. Unlike Sharma et al., who focused on a large-scale online peer-responder group, this study specifically examined pre-service counselors working with simulated cases, showing that AI feedback is also effective in more formal educational settings. These consistent patterns further support the value of GenAI models in promoting thoughtful, empathetic interactions across varied learning contexts.

In contrast to the concerns raised by participants in Gore and Dove's (2025) scoping study—who feared that LLM use might weaken genuine counseling competencies—this research revealed more balanced perspectives in post-intervention interviews. While some students initially expressed skepticism, interviewees reported feeling more comfortable as they recognized that the chatbot's feedback did not replace human learning but instead encouraged self-correction. Gore and Dove highlighted worries about limited personal growth and reduced authenticity; the feedback in this study helped ease such concerns by promoting reflective practice and reinforcing the direct application of course content. Rather than discouraging deep engagement with learning resources, the AI-based approach appeared to strengthen them, suggesting that well-structured chatbot platforms may avoid the pitfalls previously anticipated. While acknowledging potential risks remains important, the findings here suggest that carefully supervised chatbot-supported training can foster both skill development and ethical reflection.

Some theoretical perspectives may help explain the observed progression. Experiential learning theory (Kolb, 1984) suggests that active practice combined with post-action feedback, reflection, and corrected implementation can accelerate the mastery of complex skills. Additionally, social cognitive theory (Bandura, 1991) holds that observing modeled behavior (even in avatar form) and refining skills through feedback loops enhances self-efficacy, which supports improved performance. These principles have been supported by empirical studies (Bryan & Brooks, 2025; Chen et al., 2024b; Hasan et al., 2024). Finally, the repeated, low-stakes

environment allowed learners to experiment with different intervention strategies without fear of harming a real person, likely reducing performance anxiety. The conversational agent's structured prompts also highlighted key skills, encouraging participants to apply them more consistently in later sessions. Together, these mechanisms likely contributed to the strong gains observed in the post-test evaluations.

Implications

The findings from this investigation inform several practical and theoretical considerations important to educational counseling training. Primarily, the positive outcomes support the idea that ongoing virtual role-plays, paired with systematic feedback – even when provided by non-human sources – can accelerate the development of competencies in complex interpersonal areas, including those essential for supporting student populations. This principle may also apply to other fields that require advanced communication and empathy skills (Yussupova & Tarman, 2025).

Additionally, the interactive sessions may foster lasting confidence and self-awareness in future counselors, through repeated practice with varied scenarios and immediate feedback. Stakeholders in counselor education may view these results as evidence that AI-driven exercises can fill an important gap, providing frequent practice opportunities that complement core instruction. Practically, these findings suggest that educational institutions can integrate chatbot-mediated simulations into existing curricula to enhance skill acquisition without needing substantial extra resources. For example, program directors could offer these simulations as weekly practice modules, giving students the chance to engage in low-stakes role-plays outside traditional class hours. This approach could be especially useful for programs with limited access to human supervisors or role-play partners, helping to address logistical challenges while maintaining high training standards. Furthermore, the scalability of chatbot systems allows institutions to implement them across various educational settings, from large universities to smaller training programs, broadening access to high-quality practice opportunities.

The results also carry implications for professional development beyond initial training. Practicing counselors could use similar AI-driven platforms for continuing education, refining their skills as therapeutic techniques and client needs evolve. For example, chatbot simulations could be customized to address specific challenges, such as working with diverse student populations or managing complex mental health scenarios, thus supporting ongoing learning throughout a counselor's career. From a broader perspective, these insights prompt a re-evaluation of how to best integrate advanced technologies with pedagogy, supporting skill-building that is both rigorous and nearly hands-on.

This study advances knowledge by providing intervention-based evidence on the effectiveness of GenAI-supported role-plays as a supplement to standard coursework for developing therapeutic skills – an area that has previously been shaped more by speculation than data. Additionally, the embedded design, which combined objective skill assessment with

learner feedback, offers a well-rounded view of both competence development and user acceptance. In practice, these results encourage training institutions to consider tailored LLM-powered modules as an integral part of mental health professional education, particularly for those preparing to work in educational settings. Such an integrated model could serve as a replicable framework, motivating program directors to incorporate AI-based role-play tasks into broader curricula aimed at developing skilled, empathetic professionals. For example, institutions could create standardized protocols for embedding chatbot simulations into practicum courses, ensuring consistent training outcomes. The findings also contribute to the wider discussion on GenAI's role in professional education by showing how generative dialogue systems can support skill development, not just knowledge acquisition.

Limitations

There were design-specific factors that may have shaped the study's outcomes and deserve careful consideration. First, the exclusive use of text-based interactions during the intervention may favor certain therapeutic skills (e.g., verbal reflection, questioning) while offering limited practice with non-verbal aspects of counseling. This difference between training (text) and assessment (video-recorded face-to-face) introduces a possible confound in measuring skill transfer. Second, using the ITIS as both the training framework and assessment tool may create a teaching-to-the-test effect that could inflate performance gains. While aligning training and assessment can improve validity, it also raises questions about whether the skills would generalize to other contexts with different evaluation criteria. Third, the reliance on a single AI platform (Gemini 2.0) with researcher-designed prompts limits generalizability, as the specific capabilities and limitations of this LLM may have shaped the intervention's effectiveness in ways that may not translate to other systems. In addition, the presence of external observers tasked with preventing crosstalk among participants may have heightened some students' sense of being evaluated, potentially amplifying performance. Finally, the short- to medium-term nature of the intervention did not assess the long-term authenticity of trainees' skills once applied in real-world counseling, suggesting that the relationship between AI-based practice and live counseling remains an important area for future research.

Suggestions for Practice and Further Research

Gray et al.'s (2024) findings on ChatGPT's ability to generate realistic standardized patient dialogues for prenatal counseling training suggest another valuable research direction: comparing the effectiveness of different AI systems and prompt structures for counseling skill development. Their observation that AI-generated role-plays effectively conveyed emotional content (65% of responses indicating emotions) aligns with this study's findings on empathy development, suggesting that various LLM implementations may support therapeutic skill practice. It is, therefore, advisable to use generative systems to create realistic standardized patient scripts. Building on the positive trends seen here, practice initiatives could integrate such technologies with conventional supervision, blending human expertise with chatbot support. The design could also evolve to include immersive environments, as suggested by Nofal

et al. (2025), who developed an interview training module combining virtual reality-based metaverse platforms with LLM-based models. Applying this approach to counselor education could address this study's limitation of text-only interactions by offering multimodal practice opportunities. Such systems could simulate not only verbal exchanges but also non-verbal cues and environmental elements, potentially enhancing the ecological validity of GenAI-supported educational psychology training.

CONCLUSION

The proposed chatbot-empowered intervention demonstrated substantial benefits for specific fundamental skills, particularly the use of feedback/summaries, guided discovery, focusing on key cognitions and behaviors, and strategy for change, while having limited or no impact on other skills. This exploration provided practice-based evidence that AI-driven simulated role-plays can produce meaningful gains in the core skills of educational counseling students. The study is among the first in this field to show that combining regular chatbot-mediated sessions with conventional instructional frameworks can effectively enhance empathic understanding, skillful questioning, and reflective feedback abilities. By incorporating both objective scores and subjective qualitative accounts, the investigation offers a comprehensive view of how digital role-plays can support learning. Overall, the results highlight the potential of using modern digital tools to strengthen counseling skills and foster active engagement with counseling pedagogy.

REFERENCES

Adipat, S. (2025). Challenging conventions: ChatGPT's controversial impacts on educational and English language teaching practices. *Educational Process: International Journal, 14*, e2025066. https://doi.org/10.22521/edupij.2025.14.66

Akhmetov, Z., Seitova, S., Kalzhanova, G., Zhiembaev, Z., & Koishybekova, A. (2025). Effective use of the tools of probability theory and mathematical statistics in solving the problems of the professional orientation of physicists and mathematicians: an interdisciplinary approach. *Pedagogies*. Advance online publication. https://doi.org/10.1080/1554480X.2024.2422609

Alotaibi, S. M. F. (2025). Determinants of Generative Artificial Intelligence (GenAI) adoption among university students and its impact on academic performance: The mediating role of trust in technology. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2025.2492785

Alshamy, A., Al-Harthi, A. S. A., & Abdullah, S. (2025). Perceptions of generative AI tools in higher education: Insights from students and academics at Sultan Qaboos University. *Education Sciences, 15*(4), 501. https://doi.org/10.3390/educsci15040501

Arnout, B. A., & Alshehri, S. M. (2025). Causal relationships between the use of AI, therapeutic alliance, and job engagement among psychological service practitioners. *Behavioral Sciences, 15*(1), 21. https://doi.org/10.3390/bs15010021

Baker, E., & Jenney, A. (2023). Virtual simulations to train social workers for competency-based learning: A scoping review. *Journal of Social Work Education, 59*(1), 8–31. https://doi.org/10.1080/10437797.2022.2039819

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Barker, L. A., Moore, J. D., & Cook, H. A. (2024). Generative artificial intelligence as a tool for teaching communication in nutrition and dietetics education – a novel education innovation. *Nutrients, 16*(7), 914. https://doi.org/10.3390/nu16070914

Baytak, A. (2024). The Content Analysis of the Lesson Plans Created by ChatGPT and Google Gemini. *Research in Social Sciences and Technology, 9*(1), 329-350. https://doi.org/10.46303/ressat.2024.19

Boyle, K., Deisenhofer, A. K., Rubel, J. A., Bennemann, B., Weinmann-Lutz, B., & Lutz, W. (2019). Assessing treatment integrity in personalized CBT: The inventory of therapeutic interventions and skills. *Cognitive Behaviour Therapy, 49*(3), 210–227. https://doi.org/10.1080/16506073.2019.1625945

Bryan, S., & Brooks, E. E. (2025). Zombie apocalypse simulation: Elevating mental health nursing education. *Teaching and Learning in Nursing*. Advance online publication. https://doi.org/10.1016/j.teln.2024.12.009

Chan, C. K. Y. (2025). AI as the therapist: Student insights on the challenges of using generative AI for school mental health frameworks. *Behavioral Sciences, 15*(3), 287. https://doi.org/10.3390/bs15030287

Chen, A., Jia, J., Li, Y., & Fu, L. (2025). Investigating the effect of role-play activity with GenAI agent on EFL students' speaking performance. *Journal of Educational Computing Research, 63*(1), 99-125. https://doi.org/10.1177/07356331241299058

Chen, D., Liu, Y., Guo, Y., & Zhang, Y. (2024a). The revolution of generative artificial intelligence in psychology: The interweaving of behavior, consciousness, and ethics. *Acta Psychologica, 251*, 104593. https://doi.org/10.1016/j.actpsy.2024.104593

Chen, P., Yang, D., Lavonen, J., Metwally, A. H. S., & Tang, X. (2024b). How do students of different self-efficacy regulate learning in collaborative design activities? An epistemic network analysis approach. *Frontiers in Psychology, 15*, 1398729. https://doi.org/10.3389/fpsyg.2024.1398729

Coohey, C., Landsman, M. J., & Neblett, K. (2025). Evaluation of a suicide counseling skills course for social work students. *Journal of Social Work Education, 61*(1), 73–87. https://doi.org/10.1080/10437797.2024.2431560

Daniel, K., Msambwa, M. M. & Wen, Z. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education, 60*(1), e70036. https://doi.org/10.1111/ejed.70036

Darmawansah, D., Hwang, G., Lin, C., & Febiyani, F. (2025). An artificial intelligence-supported GFCA learning model to enhance L2 students' role-play performance, English speaking and interaction mindset. *Educational Technology Research and Development*. Advance online publication. https://doi.org/10.1007/s11423-025-10453-6

El Khayati, M., Courr, A., Kich, I., & Hibbi, F.-Z. (2025). Integrating artificial intelligence into high-school computer science curriculum: A perspective study in Morocco. *Journal Of Curriculum Studies Research, 7*(1), 26-46. https://doi.org/10.46303/jcsr.2025.2

Eryilmaz, A., & Başal, A. (2024). Rational AIs with emotional deficits: ChatGPT vs. counselors in providing emotional reflections. *Current Psychology, 43*, 34962–34977. https://doi.org/10.1007/s12144-024-06947-w

Fung, L., & Laing, R. (2024). A proof of concept study on the use of large language models as a client in typed role plays for training therapists. *Discover Psychology, 4*, 201. https://doi.org/10.1007/s44202-024-00302-7

Gore, S., & Dove, E. (2025). Ethical considerations in the use of artificial intelligence in counselling and psychotherapy training: A student stakeholder perspective – A pilot study. *Counselling and Psychotherapy Research, 25*, e12770. https://doi.org/10.1002/capr.12770

Gray, M., Baird, A., Sawyer, T., James, J., DeBroux, T., Bartlett, M., Krick, J., & Umoren, R. (2024). Increasing realism and variety of virtual patient dialogues for prenatal counseling education through a novel application of ChatGPT: Exploratory observational study. *JMIR Medical Education, 10*, e50705. https://doi.org/10.2196/50705

Green, G., & Adawi, S. (2024). Understanding clinical learning quality aspects in mental health nursing practice course among students: A comprehensive examination. *Nursing Reports, 14*(2), 1136–1147. https://doi.org/10.3390/nursrep14020086

Großmann, L., & Krüger, D. (2024). Assessing the quality of science teachers' lesson plans: Evaluation and application of a novel instrument. *Science Education, 108*, 153–189. https://doi.org/10.1002/sce.21832

Grunhaus, C., & Martin, C. (2025). Incorporating improv theater into online counselor education. *Journal of Creativity in Mental Health*. Advance online publication. https://doi.org/10.1080/15401383.2024.2369058

Haber, Y., Shoval, D. H., Levkovich, I., Yinon, D., Gigi, K., Pen, O., Angert, T., & Elyoseph, Z. (2025). The externalization of internal experiences in psychotherapy through generative artificial intelligence: A theoretical, clinical, and ethical analysis. *Frontiers in Digital Health, 7*, 1512273. https://doi.org/10.3389/fdgth.2025.1512273

Hamed, A., Moralejo, D., & Durante, A. (2025). Challenges and considerations in naming true and quasi-experimental research designs: A methodological discussion. *Creative Nursing*. Advance online publication. https://doi.org/10.1177/10784535241307932

Hasan, M. M., Chang, Y., Lim, W. M., Kalam, A., & Shamim, A. (2024). A social cognitive theory of customer value co-creation behavior: Evidence from healthcare. *Journal of Health Organization and Management, 38*(9), 360–388. https://doi.org/10.1108/jhom-02-2024-0074

Hinzmann, D., Wijnen-Meijer, M., Corazza, L., Becker, V., Kagerbauer, S., Haseneder, R., Berberat, P. O., & Jedlicska, N. (2023). Experiences of medical students and nursing trainees from unexpected death through simulation training. *BMC Medical Education, 23*, 667. https://doi.org/10.1186/s12909-023-04638-x

Horton, E., & Lloyd-Hazlett, J. (2025). "A reward at the end of the tunnel": A phenomenological analysis of counselors'-in-training developmental experience with frequent role-plays. *Journal of Creativity in Mental Health, 20*(1), 127–143. https://doi.org/10.1080/15401383.2024.2342244

Ibrahim, A.H. Ajlouni, A.O. (2024). Exploring the Role of ChatGPT in Supporting Special Education Undergraduates in Achieving CEC Standards: Students' Perception of the Benefits and Challenges, *Journal of Social Studies Education Research, 15*(5), 87-119. https://jsser.org/index.php/jsser/article/view/6199/710

Karjus, A. (2025). Machine-assisted quantitizing designs: Augmenting humanities and social sciences with artificial intelligence. *Humanities and Social Sciences Communications, 12*, 277. https://doi.org/10.1057/s41599-025-04503-w

Kinder, A., Briese, F. J., Jacobs, M., Dern, N., Glodny, N., Jacobs, S., & Leßmann, S. (2025). Effects of adaptive feedback generated by a large language model: A case study in teacher education. *Computers and Education Artificial Intelligence, 8*, 100349. https://doi.org/10.1016/j.caeai.2024.100349

Koishybekova, A., Seitova, S., Zhanatbekova, N., Khaimuldanov, Y., & Orazbayeva, A. (2024). Modern methods of organising independent student work in mathematical disciplines at universities. *International Journal of Mathematical Education in Science and Technology, 56*(6), 1066–1082. https://doi.org/10.1080/0020739X.2024.2315108

Kolb, D. A. (1984). Experiential learning: Experience as the source of learning and development. Prentice Hall

Krach, S. K., & Corcoran, S. (2024). Will computers replace school psychologists? An analysis of tech-based tools for assessment, consultation, and counseling. *Contemporary School Psychology, 28*(4), 496–508. https://doi.org/10.1007/s40688-023-00455-7

Kuhail, M. A., Alturki, N., Thomas, J., Alkhalifa, A. K., & Alshardan, A. (2024). Human-human vs human-AI therapy: An empirical study. *International Journal of Human-Computer Interaction*. Advance online publication. https://doi.org/10.1080/10447318.2024.2385001

Levkovich, I., & Omar, M. (2024). Evaluating of BERT-based and large language MOD for suicide detection, prevention, and risk assessment: A systematic review. *Journal of Medical Systems, 48*, 113. https://doi.org/10.1007/s10916-024-02134-3

Luo, X., Zhang, A., Li, Y., Zhang, Z., Ying, F., Lin, R., Yang, Q., Wang, J., & Huang, G. (2024). Emergence of artificial intelligence art therapies (AIATs) in mental health care: A systematic review. *International Journal of Mental Health Nursing, 33*(6), 1743–1760. https://doi.org/10.1111/inm.13384

Maurya, R. K. (2024a). A qualitative content analysis of ChatGPT's client simulation role-play for practising counselling skills. *Counselling and Psychotherapy Research, 24*(2), 614–630. https://doi.org/10.1002/capr.12699

Maurya, R. K. (2024b). Using AI based chatbot ChatGPT for practicing counseling skills through role-play. *Journal of Creativity in Mental Health, 19*(4), 513–528. https://doi.org/10.1080/15401383.2023.2297857

Mi, Y., Rong, M., & Chen, X. (2025). Exploring the affordances and challenges of GenAI feedback in L2 writing instruction: A comparative analysis with peer feedback. *ECNU Review of Education*. Advance online publication. https://doi.org/10.1177/20965311241310883

Murtaza, M., Cheng, C., Albahlal, B. M., Muslam, M. M. A., & Raza, M. S. (2025). The impact of LLM chatbots on learning outcomes in advanced driver assistance systems education. *Scientific Reports, 15*, 7260. https://doi.org/10.1038/s41598-025-91330-3

Muryono, S. (2024). Evaluating Counseling Practices through Positive Psychology for University Students. *Journal of Social Studies Education Research, 15*(3), 118-145. https://jsser.org/index.php/jsser/article/view/5875/682

Nelson, J., Kaplan, J., Simerly, G., Nutter, N., Edson-Heussi, A., Woodham, B., & Broman-Fulks, J. (2025). The balance and integration of artificial intelligence within cognitive behavioral therapy interventions. *Current Psychology*. Advance online publication. https://doi.org/10.1007/s12144-025-07320-1

Nguyen, K. V. (2025). The use of generative AI tools in higher education: Ethical and pedagogical principles. *Journal of Academic Ethics*. Advance online publication. https://doi.org/10.1007/s10805-025-09607-1

Nofal, A. B., Ali, H., Hadi, M., Ahmad, A., Qayyum, A., Johri, A., Al-Fuqaha, A., & Qadir, J. (2025). AI-enhanced interview simulation in the metaverse: Transforming professional skills training through VR and generative conversational AI. *Computers and Education Artificial Intelligence, 8*, 100347. https://doi.org/10.1016/j.caeai.2024.100347

Prescott, J., Ogilvie, L., & Hanley, T. (2024). Student therapists' experiences of learning using a machine client: A proof-of-concept exploration of an emotionally responsive interactive client (ERIC). *Counselling and Psychotherapy Research, 24*, 524–531. https://doi.org/10.1002/capr.12685

Shaleh M., Awad F.B., Rezki A. (2022). Counseling Based on Local Wisdom for Conflict Settlement in the Workplace: A Case Study in Kendari. *Journal of Social Studies Education Research, 13* (1), pp. 125 – 142.

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2023). Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence, 5*(1), 46–57. https://doi.org/10.1038/s42256-022-00593-2

Sudirman, S., Belbase, S., Rodríguez-Nieto, C., Muslim, A., & Faizah, S. (2025). Personalization of interactive teaching materials supported by augmented reality: Potentials vs obstacles in 3D geometry learning. *Journal Of Curriculum Studies Research, 7*(1), 152-178. https://doi.org/10.46303/jcsr.2025.8

Sun, Z., Anyanwu, C., Liu, Y., & Arifani, Y. (2025). A case study of teacher-industry collaborative teaching practice in digital media art design course at Chinese higher vocational colleges. *Journal Of Curriculum Studies Research, 7*(1), 89-108. https://doi.org/10.46303/jcsr.2025.5

Symeou, L., Louca, L., Kavadella, A., Mackay, J., Danidou, Y., & Raffay, V. (2025). Development of evidence-based guidelines for the integration of generative AI in university education through a multidisciplinary, consensus-based approach. *European Journal of Dental Education, 29*(2), 285-303. https://doi.org/10.1111/eje.13069

Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education Artificial Intelligence, 8*, 100355. https://doi.org/10.1016/j.caeai.2024.100355

Terzi, A., Yildirim, Y., & Kocakoc, I. D. (2025). The power of role play in nursing education: A bibliometric analysis and visualization. *Nurse Education Today, 149*, 106633. https://doi.org/10.1016/j.nedt.2025.106633

Turarbek, A., Bektemesov, M., Ongarbayeva, A., Orazbayeva, A., Koishybekova, A., & Adetbekov, Y. (2023). Deep convolutional neural network for accurate prediction of seismic events. *International Journal of Advanced Computer Science and Applications, 14*(10), 604-613. https://doi.org/10.14569/ijacsa.2023.0141064

Vowels, L. M., Francois-Walcott, R. R., & Darwiche, J. (2024). AI in relationship counselling: Evaluating ChatGPT's therapeutic capabilities in providing relationship advice. *Computers in Human Behavior Artificial Humans, 2*(2), 100078. https://doi.org/10.1016/j.chbah.2024.100078

Yussupova, S., & Tarman, B. (2025). Cultural empathy and school climate in Kazakhstan: A factor and cluster analysis of students' perceptions. *Journal Of Curriculum Studies Research, 7*(1), 47-65. https://doi.org/10.46303/jcsr.2025.3

Yusuf, H., Money, A., & Daylamani-Zad, D. (2025). Pedagogical AI conversational agents in higher education: A conceptual framework and survey of the state of the art. *Educational Technology Research and Development*. Advance online publication. https://doi.org/10.1007/s11423-025-10447-4

Zheng, Y. B., Zhou, Y. X., Chen, X. D., & Ye, X. D. (2025). The influence of large language models as collaborative dialogue partners on EFL English oral proficiency and foreign language anxiety. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2025.2453191

APPENDIX

Case scenarios

1. Fatima, a 25-year-old Native Kirgiz student, is studying at a prominent university in Kazakhstan. She comes from a village on periphery of the Kyrgyz Republic, where she grew up immersed in her Native Kirgiz culture and traditions. Emily is pursuing a master program in nursing and wants to go back to her village after finishing the program to support her community. Emily is facing many issues such as cultural adjustment, stereotyping and misunderstanding from her peers and professors, and limited academic support. She visits the university counseling center seeking support.
2. Arman, 21 years old, a sophomore, is a victim of cyberbullying, with hurtful messages and rumors being spread online. He seeks guidance from the university counselor to address the cyberbullying, learn about online safety measures, and develop coping strategies to protect his well-being in the digital world.
3. Chingiz is 17 years old, 12th grader, a male, recently experienced the loss of a close family member, his grandmother, and is struggling with grief. He seeks assistance from the school counselor to navigate the grieving process, find healthy coping mechanisms, and access additional support resources to help him cope with his loss.
4. Erzhan is 16 years old returnee (Chinese-born with Kazakh background) male. Erzhan's parents recently divorced, and he's struggling to cope with the changes at home. He seeks guidance from the school counselor to process his emotions, develop healthy coping strategies, and find ways to maintain focus on his schoolwork despite the family difficulties.
5. Kairat is confused about his future after high school. He seeks guidance from the school counselor to explore various college and career options, identify his interests and strengths, and develop a plan for post-secondary education or vocational training that aligns with his long-term goals.
6. Aidana, a high-achieving student, is overwhelmed by her heavy course load and extracurricular activities. She seeks guidance from the school counselor to cope with her academic stress, develop effective time management skills, and strike a balance between her responsibilities and personal well-being.
7. Growing up in a rural, low-socioeconomic community, Sabira faced numerous barriers as she sat alongside her eight-grade classmates with her high school schedule in hand. The oldest of four children, Sabira was entrusted with many parental responsibilities as her single mother, with only her high school education, worked two jobs trying to support the family. Sabira lived

in an area historically divided by race issues. As a Tatar student attending a school of predominantly Kazakh students, Sabira was constantly trying to make sense of the various messages she received about her potential as a student while balancing attention issues in classes. Sabira came to the counselor with tears in her eyes saying “I cannot do this. My family needs me. My mom says I am being selfish by putting time and money into school when I should get a job and help her. I have had to work so hard in middle school. I will never be smart enough for college anyway. I don’t know why she sighed this.”

8. Juri, 24 years old Korean Kazakh, is pursuing engineering at a university. He belongs to a lower-income family background and is the first-generation learner to college. Juri has always been a good student with excellent academic performance. Being the first person in the family going to college, his family has high expectations from him which creates pressure on him. Lately, he has been experiencing intense anxiety and stress related to his studies. Juri often feels overwhelmed and is constantly worried about failing or not meeting his own high expectations. He has been isolating himself from friends and family, fearing that he will disappoint them if he doesn’t achieve top grades.

Table 1.

Basic statistics for therapeutic skills by group and time point. mean (standard deviation)

Variable	Group	Pre-test	Post-test
Pacing and efficient use of time	Control	2.71 (0.81)	2.88 (0.65)
	Role-play	3.04 (0.76)	3.19 (0.74)
Clarity of communication	Control	4.17 (0.75)	4.03 (0.76)
	Role-play	3.94 (0.77)	4.25 (0.62)
Use of feedback/summaries	Control	3.29 (0.81)	3.24 (0.93)
	Role-play	3.11 (0.95)	3.85 (0.93)
Rationale	Control	2.14 (1.07)	2.37 (0.79)
	Role-play	1.98 (0.57)	2.19 (0.62)
Guided discovery	Control	2.95 (0.95)	3.34 (0.78)
	Role-play	2.77 (1.15)	4.02 (0.91)
Therapeutic relationship/collaboration	Control	3.53 (0.68)	3.80 (0.69)
	Role-play	4.06 (0.79)	4.17 (0.70)
Handling problems/questions/objections	Control	3.22 (0.70)	3.36 (1.0)
	Role-play	3.49 (0.75)	3.64 (0.81)
Empathic understanding	Control	4.02 (0.78)	4.12 (0.93)
	Role-play	4.26 (0.90)	4.36 (0.76)
Focusing on key cognitions and behaviors	Control	3.07 (0.69)	3.17 (0.70)
	Role-play	3.28 (0.72)	3.49 (0.72)
Strategy for change	Control	2.71 (0.72)	3.07 (0.49)
	Role-play	2.91 (0.99)	3.66 (0.96)

	Control	2.10 (0.55)	2.39 (0.59)

Table 2.

Themes, definitions, and representative quotes

Theme	Sub-theme	Definition	Quote
Engagement with AI Client	Moments of Immersion	Instances where students felt genuinely engaged with the simulated client as if in a real counseling session.	"There were times, especially with the scenario about Fatima adjusting to university life, where I almost forgot it was not a real human. Well, the way she described feeling isolated really resonated."
	Navigating AI Limitations	Challenges students faced due to the nature of interacting with an AI rather than a human.	"Sometimes it felt a bit clunky. Like, I asked a question I thought was open-ended, but the response was super short, and I was like, 'Okay, what is on?' That was tough."
Impact of AI Tutor	Value of Immediate Feedback	Students' appreciation for the instant performance review provided after each session.	"The tutor bot popping up right after with feedback was gold. It pointed out when I maybe jumped to conclusions too quickly instead of just listening."
	Feedback Driving Reflection	How the tutor's comments prompted students to think more deeply about their counseling techniques.	"Getting that feedback made me really think about why I asked certain questions. Like, was I actually trying to understand them better, or was I just trying to move the conversation along?"
Skill Development	Safe Space for Practice	The perception of the simulation as a low-stakes environment to experiment with counseling skills.	"It was great being able to practice without the burden of having a real person in front of you, worrying if you are going to say the wrong thing and mess them up."
	Recognizing Skill Application	Students becoming more aware of when and how they were utilizing specific counseling techniques.	"I noticed myself consciously trying to reflect back what that person was saying more often, and the tutor bot usually picked up on that, which was encouraging."

Emotional Responses	Moments of Frustration	Instances where the interaction led to feelings of annoyance or difficulty.	"With the cyberbullying case, Arman's responses felt a bit repetitive at one point, and I got a bit frustrated trying to find a new angle to explore it with him."
	Experiences of Empathy	Moments when students felt an emotional connection or understanding towards the simulated client's situation.	"Sabira's story about family pressure and how she doubted herself… even though I knew it was a simulation, it tugged at my heartstrings a bit. I felt the weight of the situation."
Lingering Impact	Post-Session Reflection	Thoughts and considerations about the simulations that stayed with students after the sessions ended.	"I found myself thinking about some of the cases afterwards, like Chingiz dealing with grief. It made me consider how I would handle similar situations if they came up with real students."
Perceived Overall Value	Preparation for Real-World Practice	Students' belief that the simulation experience contributed to their readiness for actual counseling scenarios.	"Honestly, I feel a bit more confident now. Having gone through different scenarios, even simulated ones, makes the idea of facing them in real life a little less daunting."