Journal of Teaching and Learning with Technology, Vol. 14, Special Issue, pp.122-145.
doi: 10.14434/jotlt.v12i1.41694

AI-Powered Simulation: Leveraging Technology to Advance Nursing Communication Skills with AI SimBot

Tiffany Kim
Northeastern University
t.kim@northeastern.edu

Yash Gopalji Pankhania
Northeastern University
pankhania.y@northeastern.edu

Abstract: Nursing students need opportunities to practice patient communication skills in academic settings before clinical practice. Standardized-patient simulations are effective but limited by high costs and logistical barriers. These constraints reduce the frequency and diversity of simulation experiences. This study evaluated the feasibility of AI SimBot, a generative artificial intelligence (GenAI) virtual patient simulation. Guided by Bandura's social cognitive theory, we examined whether AI SimBot enhanced students' confidence and readiness to engage in sensitive communication with adolescents. AI SimBot was implemented in a prelicensure public/community health nursing course. Ninety-eight students practiced administering the CRAFFT screening tool with a virtual adolescent patient, "Jordan," received rubric-based feedback, and submitted reflections. A qualitative descriptive design was used. Analyses focused on two reflection prompts (confidence and authenticity), with three coders independently coding and reaching consensus on themes. Students initially described nervousness but reported marked increases in confidence following the simulation. Themes highlighted the CRAFFT tool as a useful guide, the importance of nonjudgmental approaches, and challenges in adolescent communication. Students found Jordan realistic, citing adolescent language and gradual disclosure, but noted limitations such as absent nonverbal cues, an unrealistically cooperative patient, and occasional technical latency. AI SimBot enhanced nursing students' perceived confidence and was experienced as an authentic and educationally valuable simulation. GenAI-based virtual patients offer a scalable, cost-effective supplement to standardized-patient encounters, contributing to competency-based education by providing repeated, feedback-rich practice in sensitive, high-stakes communication.

Keywords: generative artificial intelligence; virtual-patient simulation; nursing education; therapeutic communication; substance-use screening; confidence

Communication skills are a foundational competency in health professions, serving as the critical bridge between clinical knowledge and quality patient care (Kwame & Petrucka, 2021). For students, developing effective therapeutic communication is essential before implementing these skills in clinical settings. While traditional simulation methods using standardized patient (SP) actors provide valuable practice opportunities, they face significant barriers, including high costs and logistical challenges, that often restrict the frequency and depth of student experiences.

The AI SimBot project emerged as a response to these constraints, leveraging generative artificial intelligence (GenAI) technology to create a virtual-patient simulation tool. AI SimBot is a generative pretrained transformer (GPT)-based simulation that enables nursing students to engage in realistic audio and text-based conversations with a virtual SP. This GenAI application allows students to practice sensitive communication skills, such as substance-use screening with adolescents, in a safe, accessible environment.

Unlike traditional simulations that depend on actor and faculty availability, AI SimBot offers unlimited opportunities for practice and provides immediate, structured feedback based on faculty-developed rubrics. By harnessing the capabilities of GenAI to create reproducible, standardized experiences, AI SimBot addresses a critical gap in clinical education while complementing traditional simulation approaches. The flexibility and scalability of this technology make it particularly valuable in addressing the practical constraints that have historically limited simulation opportunities in nursing education.

This case study examined the development, implementation, and preliminary evaluation of AI SimBot within a prelicensure public/community health nursing course. Guided by Bandura's (1977) social cognitive theory, which highlights the role of self-efficacy in shaping learners' willingness to engage in challenging tasks, the study focused on two core research questions (RQs):

- RQ1: How does participation in AI SimBot influence nursing students' self-perceived confidence in discussing substance use with adolescents?
- RQ2: How do students perceive the authenticity and educational value of the AI virtual patient, and what improvements do they recommend for future iterations?

Literature Review

This review synthesizes scholarly literature across three domains essential to understanding AI SimBot's development: communication simulation in clinical education, substance-use screening pedagogy, and GenAI in virtual simulation.

Communication Simulation in Nursing Education

Effective communication skills are a core competency for all health care professionals, essential for developing therapeutic relationships and providing high-quality patient care across clinical settings (Kwame & Petrucka, 2021). This fundamental skill enables nurses to understand patient needs, provide education, offer emotional support, and collaborate with interprofessional teams (Blackmore et al., 2018; Webster, 2014). However, providing students with opportunities to practice therapeutic communication with actual patients presents challenges in many clinical environments due to high patient acuity, shorter lengths of stay, limited clinical placements, and concerns about patient safety and comfort (Bloomfield et al., 2015; Webster, 2014).

Simulation involving SPs has emerged as an effective method for teaching therapeutic communication skills across nursing education and other health professions (Becker et al., 2006; Blackmore et al., 2018). SPs (individual actors trained to present an illness or scenario in a standardized manner) allow students to practice communication skills in a controlled, safe environment without risk to actual patients. After participating in SP exercises, nursing students show significant gains in communication criteria including active listening and appropriate limit setting (Webster, 2014). In a systematic review, Blackmore et al. (2018) found that most professional–patient communication studies relied on SPs, with generally positive effects on learner satisfaction and knowledge. However, the authors noted only limited evidence of translation to real-patient outcomes, a gap echoed by Rutherford-Hemming et al. (2019), who also called for stronger study designs and patient-oriented measures. Gutiérrez-Puertas et al. (2020), reviewing 19 communication-simulation interventions, concluded that although most improved student performance, the majority relied on quasiexperimental designs, self-report instruments, and short-term follow-up, emphasizing the need for scalable, theory-anchored approaches with objective assessment.

A significant body of research has demonstrated that simulation experiences enhance not only skill development but also students' confidence and self-efficacy in their communication abilities (Al Gharibi et al., 2021; Al Khasawneh et al., 2021; Foucault-Fruchard et al., 2024; Hsu et al, 2015; Jallad, 2024; Yun et al., 2023). Self-efficacy, defined as an individual's belief in their capability to perform specific tasks successfully, is particularly important in nursing education as it influences students' willingness to engage in therapeutic communication with patients and health care team members. Bandura's (1977) social cognitive theory provides a theoretical framework for understanding how simulation develops self-efficacy through four primary sources: mastery experiences, vicarious learning, verbal persuasion, and emotional states. In simulation-based education, mastery experiences occur when learners successfully practice a skill, such as administering a screening tool. Vicarious learning is supported when students observe peers or review modeled scenarios. Verbal persuasion is enacted through structured feedback and encouragement, reinforcing learners' belief in their competence. Emotional states also play a role, as psychologically safe environments can reduce anxiety and help students regulate stress while rehearsing challenging tasks.

Multiple studies have reported significant improvements in students' communication self-efficacy following simulation experiences. In a randomized controlled trial, Hsu et al. (2015) found that nurses in scenario-based simulation training demonstrated significantly greater communication self-efficacy compared to those in traditional classroom-based courses. Participants in the simulation group also reported higher satisfaction with the learning experience. Ok et al. (2020) found that nursing students who participated in an SP simulation focused on therapeutic communication prior to their psychiatric clinical rotation demonstrated reduced anxiety and improved communication skills in subsequent clinical practice. Interprofessional simulations also show promise with both MacLeod et al. (2022) and Granheim et al. (2018) finding that team-based scenarios improved nursing students' confidence in interprofessional communication.

Researchers have cautioned against relying solely on self-reported confidence measures. Blackmore et al. (2018) identified discrepancies between perceived communication self-efficacy and objectively measured performance, highlighting the importance of using multiple assessment methods. Despite this caveat, evidence strongly supports simulation as an effective educational strategy for building communication confidence essential for establishing therapeutic relationships and functioning effectively within health care teams.

Substance-Use Screening Pedagogy

Substance use disorders (SUDs) present a critical public health challenge in the United States. In 2023, an estimated 48.5 million individuals aged 12 years or older (17.1% of the U.S. population) met SUD criteria, yet only 14.6% received any form of treatment, leaving more than 41 million without care (Substance Abuse and Mental Health Services Administration [SAMHSA], 2024b). Among adolescents aged 12 to 17 years, approximately 2.2 million individuals (8.5%) met SUD criteria in the past year, highlighting the need for early intervention (SAMHSA, 2024b). Early substance use is linked to lasting cognitive, behavioral, and health consequences (Mackavey & Kearney, 2020), reinforcing the urgency of equipping health professionals with skills in early detection and intervention.

Nurses are uniquely positioned to serve as a first line of defense in addressing substance misuse, especially in nonspecialty settings such as primary care and school-based health clinics (O'Brien et al., 2019). Despite their frontline role, only 0.1% of nurses reported specialization in substance use care, reflecting a significant training and workforce gap (O'Brien et al., 2019). Their foundational expertise in patient education, health promotion, and communication make nurses

ideal providers of preventive interventions. Nonetheless, substance-use screening remains infrequent in practice. National recommendations endorse universal screening in primary care (SAMHSA, 2020; U.S. Preventive Services Task Force, 2020), yet research has indicated that only about one-third of adults and half of adolescents receive such screening (Gainey et al., 2022). Contributing factors include provider discomfort, unfamiliarity with screening tools, and limited training (Adamshick & Payton, 2024; Mackavey & Kearney, 2020).

Screening, brief intervention, and referral to treatment (SBIRT) is an evidence-based, public health framework designed to identify and intervene with individuals at risk for or experiencing substance misuse. The model is structured around motivational interviewing (MI) techniques, which emphasize empathy, patient autonomy, and collaborative problem solving. SBIRT has demonstrated efficacy in reducing risky substance use and is endorsed by major federal health agencies (Koetting & Freed, 2017; SAMHSA, 2024a). Its integration into nursing education addresses procedural training and promotes development of therapeutic communication, clinical judgment, and patient-centered care.

The CRAFFT test is a screening instrument developed specifically for adolescents—a simple, verbally administered tool whose name is a mnemonic for its six questions (car, relax, alone, forget, friends, trouble). The CRAFFT test directly operationalizes SBIRT by establishing a clear trigger for further action. It is valid and reliable for screening adolescents for substance-related problems, with good sensitivity and specificity (Knight et al., 2002) and is recommended by American Academy of Pediatrics practice guidelines (Hagan et al., 2017).

Educational strategies for teaching SBIRT increasingly highlight the importance of therapeutic communication and clinical judgment. MI is central to this effort, as it fosters trust, encourages honest disclosure, and supports patient-centered decision making. However, many students report limited confidence in applying MI techniques due to a lack of practice opportunities (O'Brien et al., 2019). Nursing programs have responded by embedding MI into SBIRT instruction, often using interactive teaching methods such as role-play and simulation (Koetting & Freed, 2017; Seigart et al., 2018). These modalities allow students to practice phrasing sensitive questions, responding to ambivalence, and transitioning between screening and intervention in real time. Debriefing sessions following simulation promote reflective learning and targeted feedback, reinforcing communication strategies and enhancing clinical reasoning (Adamshick & Payton, 2024; Hitchcock et al., 2019).

SBIRT education appears to improve students' attitudes toward patients who use substances and increase their sense of professional responsibility. Students have frequently reported feeling more confident, prepared, and "role secure" after training, a shift that is essential to reducing stigma and promoting routine screening (Adamshick & Payton, 2024; Gainey et al., 2022). Knopf-Amelung et al. (2018) found that active learning approaches not only improved SBIRT competence but also fostered the belief that substance-use screening is a fundamental nursing duty.

GenAI in Virtual-Patient Simulation

GenAI, including large language models (LLMs), chatbots, and AI-driven virtual patients, is rapidly emerging as a transformative tool in health professions education. These technologies promise to supplement traditional teaching methods by simulating patient encounters in a flexible, on-demand manner, addressing significant barriers in conventional communication skills training, which often relies on human SP actors, that can be costly, labor intensive, and stressful for learners, with limited opportunities for repetition (Cook et al., 2025; Holderried et al., 2024). Educators worldwide have begun exploring GenAI platforms to support clinical communication training, accelerated by advancements in LLMs such as GPT, which have demonstrated remarkable capabilities in

generating human-like text and maintaining context-aware conversations (Karabacak et al., 2023; Montejo et al., 2024; Öncü et al., 2025).

One major advantage of AI-powered virtual patients is their ability to provide consistent, repeatable practice opportunities without the logistical challenges of scheduling SPs. In a study using a GPT-3.5-powered chatbot for medical students practicing history taking, researchers found that 97.9% of AI answers were medically plausible, with positive user feedback (~77/100 usability) that allowed repetitive practice without live evaluator pressure (Holderried et al., 2024). Cook et al. (2025) demonstrated that LLM-based virtual patients could simulate authentic, preference-sensitive dialogues while delivering personalized performance feedback, with GPT-4.0-Turbo feedback rated comparably to human expert feedback, highlighting the model's potential not only to replicate patient behavior but also to serve as a formative assessment tool for clinical educators.

Emerging studies have underscored the broader educational value of GenAI in enhancing learners' clinical competence across contexts. Brügge et al. (2024) showed that students engaging in repeated LLM-based patient interviews with structured feedback demonstrated statistically significant gains in clinical decision-making scores, particularly in contextualization and information gathering. García-Torres et al. (2024), in a systematic review, concluded that AI-enhanced virtual patients (especially those with natural-language-processing capabilities) consistently improved student satisfaction and clinical reasoning performance. Sardesai et al. (2024) and Benfatah et al. (2024) further highlighted the accessibility and flexibility of GenAI tools, noting high learner engagement and usability. Taken together, these studies support the growing consensus that GenAI platforms can offer scalable, pedagogically sound alternatives to human actors, particularly when enhanced with structured feedback, contextual variation, and targeted learning objectives.

Beyond basic communication practice, GenAI offers unique capabilities for personalized feedback and adaptive learning. In a randomized controlled trial, students who received AI-generated feedback after virtual-patient encounters showed significantly greater improvement in clinical reasoning performance compared to those without feedback (Brügge et al., 2024). AI-driven tools can also be programmed to portray various case scenarios and demographic characteristics, offering exposure to diverse clinical situations difficult to achieve with standard simulations (Cook et al., 2025).

While studies show considerable promise, research has also identified important limitations. One recurrent concern is potential inaccuracies or "hallucinations" in AI responses; Holderried et al. (2024) found approximately 2% of AI-simulated patient responses were implausible. Additional challenges include disconnection issues and language-processing errors (Öncü et al., 2025). LLMs may exhibit unnatural conversational patterns (excessive agreeableness, verbose phrasing, or failure to challenge inappropriate clinical decisions), compromising authenticity (Cook et al., 2025). These systems also lack capacity to convey complex emotional states or nuanced nonverbal cues (Benfatah et al., 2024; García-Torres et al., 2024). Because of these limitations, most educators emphasize that AI-powered virtual patients should supplement, not replace, SP programs or real clinical encounters (Sardesai et al., 2024). When implemented thoughtfully and with attention to their current limitations, GenAI tools can offer meaningful opportunities for repetitive, low-stakes practice that supports learners' transition from classroom knowledge to clinical reasoning.

Summary

Collectively, the literature underscores three converging findings. First, repeated, high-fidelity practice with feedback is indispensable for mastering therapeutic communication, yet logistical constraints limit students' exposure to SP simulations. Second, effective substance-use screening education must integrate SBIRT procedural skills with MI-driven communication practice, but such

integrated training opportunities are often absent. Third, recent advances in GenAI show promise for delivering scalable, conversationally rich virtual-patient encounters, though evidence on their application to substance-use communication is still emerging.

A clear research gap exists for a single, scalable intervention addressing all three issues simultaneously. Specifically, there is limited research on educational tools providing unlimited practice in integrated CRAFFT and MI training that leverage AI to deliver automated feedback. These intersecting gaps led to AI SimBot's development, which leveraged LLM technology to give nursing students unlimited practice sessions in adolescent substance-use screening, each with immediate, structured feedback, thereby addressing documented barriers while building on the pedagogical strengths identified across all three domains.

Method

Building AI SimBot: System Design and Functionality

The AI SimBot application is built on a multi-layered architecture using Streamlit, a Python-based web framework chosen for its scalability and ease of deployment. The system integrates a suite of OpenAI models to simulate a realistic patient interaction: GPT-4o serves as the core conversational engine, OpenAI Whisper provides speech-to-text transcription, and the TTS-1 model handles text-to-speech synthesis.

The conversational flow for each turn of dialogue involves five stages. First, the Streamlit interface records the student's speech. This audio is then transcribed into text by the Whisper model, which is adept at handling diverse accents and background noise (FreeCodeCamp, 2024). The resulting text is sent to the GPT-4o model, which generates a response from the perspective of the patient, "Jordan." Jordan's textual response is then converted into high-quality spoken audio by the TTS-1 model, using the "nova" voice to align with the patient's profile and ensuring a rapid response time for natural conversational pacing. Finally, the Streamlit interface plays the audio response for the student.

From the user's perspective, the interface guides the student through the simulation. An introductory page provides access to login and instructional materials. After logging in and completing a microphone and speaker check, the student enters the main chat interface to interact with the virtual patient. An option to toggle between voice and text-based chat is available to accommodate any technical difficulties.

After the session ends, the system provides two key outputs: rubric-based feedback and a complete downloadable transcript. For the feedback function, the AI transitions from the role of patient to that of an evaluator. While GPT-4o excelled as the patient, it generated inaccurate and overly positive (sycophantic) feedback to students. Therefore, the feedback mechanism was updated to use OpenAI's O3-mini model, which delivered significantly more accurate and robust evaluations based on the faculty-developed rubric. The switch to a reasoning model was made to ensure the feedback was accurately grounded in the rubric and the student's encounter with Jordan; although more powerful models were available, O3-mini was sufficient for this evaluative task while being faster and more cost effective than other options at the time.

The system also generates a transcript of the entire exchange, formatted with clear speaker identification for easy review. This downloadable document serves multiple purposes: It enables student self-reflection, provides objective data for faculty assessment, and can be used for quality improvement research. To facilitate iterative development and instructional oversight, the application includes an administrative interface with role-based access controls. This backend allows faculty to modify the AI's prompts, preview its behavior, and review chat histories, which is essential

for refining the simulation scenarios without starting each test conversation from the beginning. The complete prompt defining Jordan's persona, the clinical scenario, and the evaluative rubric for this study is available in Appendix 1. To see AI SimBot in action or leverage the open-source code to develop similar virtual-patient simulations, a video demonstration and the project repository are also available at https://aisimbot.github.io/.

The Classroom Setting

AI SimBot was integrated as a mandatory assignment into a prelicensure public/community health nursing course at Northeastern University. Ninety-eight students across two campuses completed didactic coursework on SUD, the use of a substance-use screening tool (CRAFFT), and harm reduction practices (general therapeutic communication and MI techniques are covered earlier in the curriculum). Using their personal devices, they then practiced administering the CRAFFT tool with AI SimBot, which role-played a 17-year-old nonbinary student named Jordan, visiting their primary care provider for a routine sports physical. After the simulation, AI SimBot provided specific feedback directly to students using a faculty-built rubric. Students submitted transcripts of their conversations with Jordan and written reflections for faculty review and grading. All students completed the assignment; while many practiced multiple times before submitting their preferred session, no students opted out or had incomplete submissions for grading.

Procedure

This study was submitted to the Institutional Review Board at Northeastern University and determined not to be human subjects research as it involved retrospective analysis of existing deidentified educational assessment data collected as part of normal educational practices. Following course completion, a designated data steward removed all student identifiers from the reflection submissions to maintain confidentiality in accordance with Northeastern University institutional policies for educational assessment activities. Students completed five brief reflections after the simulation; consistent with the study's feasibility aims and theoretical framing on self-efficacy and perceived fidelity, the primary author and two graduate nursing students analyzed Prompts 1 (pre/post confidence) and 5 (authenticity/realism). The remaining prompts were retained for context but are not thematically reported in this article. Using a qualitative descriptive approach, the coders independently coded all responses and then met to reconcile codes and agree on themes by consensus (no kappa calculated).

Results

Findings are presented in relation to the two research questions. RQ1 examined how participation in AI SimBot influenced nursing students' self-perceived confidence in discussing substance use with adolescents. RQ2 explored students' perceptions of the authenticity and educational value of the virtual patient.

RQ1: How Does Participation in AI SimBot Influence Nursing Students' Self-Perceived Confidence in Discussing Substance Use With Adolescents?

Analysis of student reflections to Prompt 1 (reflect on your confidence discussing substance use with adolescent patients before and after the virtual simulation) revealed five themes related to their confidence discussing substance use with adolescent patients before and after participating in the

virtual simulation: (1) initial nervousness and uncertainty, (2) increased confidence postsimulation, (3) the CRAFFT tool as a guide, (4) importance of a nonjudgmental approach, and (5) challenges of adolescent communication. These themes illustrate students' evolving confidence and insights into effective communication strategies with adolescents.

Initial Nervousness and Uncertainty

Prior to engaging in the virtual simulation, many students (n = 49; 50% frequency) expressed feeling nervous, hesitant, or uncertain about discussing substance use with adolescents. This anxiety stemmed from inexperience with adolescent patients, concern about approaching sensitive topics appropriately, and worry about potential negative reactions from patients. Students noted: "Before the virtual simulation, I felt a bit nervous about how to approach questions related to substance use, especially since I had never interacted with an adolescent in this context before," and "I was concerned about coming across as judgmental or making the patient uncomfortable." These concerns reflected students' awareness of the delicate nature of substance use discussions and the importance of creating a safe communication environment.

Increased Confidence Postsimulation

Following the virtual simulation experience, most students (n = 69; 70% frequency) reported significant improvements in their confidence levels regarding substance use discussions with adolescents. This transformation was often attributed to the practical experience gained. Students stated: "After completing the simulation, I feel much more confident in my ability to navigate substance use screenings," and "The experience allowed me to practice my communication skills in a safe environment, learn how to ask open-ended questions, and explore how to create a supportive atmosphere for adolescents." Others noted: "I even did it twice and felt more prepared for the real-life situation with more practice," and "This definitely helped with my confidence. As someone who has never drank or taken any drugs, this aspect of health care has been difficult for me to navigate as the terminology can be very new for me."

The CRAFFT Tool as a Guide

Students (n = 29; 30% frequency) emphasized the important role that the structured CRAFFT screening tool played in building their confidence and guiding their approach to substance-use discussions. Students reflected: "The structured nature of the CRAFFT tool provided a clear framework, which made it easier to navigate the conversation while ensuring all necessary topics were covered," and "The experience reinforced how important it is to actively listen, maintain a calm tone, and use validated screening tools like CRAFFT to guide the conversation effectively."

Importance of a Nonjudgmental Approach

Students (n = 51; 52% frequency) consistently identified the significance of maintaining a nonjudgmental communication approach when discussing substance use with adolescent patients. Through the simulation, they recognized that creating a safe, supportive environment was essential for encouraging honest disclosure. One student explained, "However, I focused on creating a welcoming and safe environment while ensuring my approach remained non-judgmental. This helped establish trust and allowed for a smooth, open conversation." Students identified specific

techniques for maintaining this nonjudgmental stance: "After the simulation, I feel much more confident in my ability to ask sensitive questions while maintaining a supportive and non-judgmental tone." Explicit reassurance was also noted as important, with one student sharing, "I tried to emphasize that these questions are standard to all patients, that I wanted Jordan to be honest with me, and that there was no judgement about any answers they provided." Empathy's role in rapport building also emerged: "I learned the value of approaching the topic with empathy and curiosity, rather than making the conversation feel condescending or intrusive." This theme underscores students' growing awareness of how their communication approach directly influences patient disclosure and engagement.

Challenges of Adolescent Communication

Students (n = 27; 28% frequency) identified specific challenges related to communicating with adolescents about sensitive topics such as substance use. One student observed the potential difficulties in establishing open communication: "Since the patient is at an age that can be sensitive and secretive, it is difficult to construct responses that will be viewed positively. It can also be difficult to have an adolescent patient open up about substance abuse." Students expressed concern about adolescents' potential reluctance to disclose information due to fear of consequences: "I know that the main worries that push adolescents to lie are fear of parents finding out and worry of judgement." Some students identified specific communication challenges that might arise during these conversations: "I was concerned about how to phrase questions effectively, respond to potential defensiveness, and ask follow-up questions if they answered yes." The complex social dynamics of adolescence were also recognized as a potential barrier to effective communication, with one student noting: "I feel like they may not be as open to the idea of change because the desire to fit in can be powerful." As another student explained: "I feel anxious about it in general because I'm concerned about the potentially negative reaction that the adolescent could have." This recognition of potential challenges indicated students' growing clinical reasoning abilities and their understanding of the need to adapt communication strategies for different patient populations.

RQ2: How Authentic Did Students Find the AI SimBot Virtual-Patient Interaction?

Reflections on Prompt 5 (How realistic did you find the virtual-patient interactions in terms of mimicking real-life clinical encounters? Were there aspects of the interaction that felt particularly authentic or inauthentic? How did this affect your learning experience?) highlighted five themes related to the realism and utility of the AI-powered virtual patient: (1) realistic adolescent language and communication style, (2) authentic hesitation and gradual disclosure, (3) lack of nonverbal communication, (4) unrealistic level of cooperation, and (5) technical limitations affecting realism.

Realistic Adolescent Language and Communication Style

Students (n = 30; 31% frequency) found the virtual patient's communication style authentic, noting that age-appropriate slang and casual language enhanced realism. Students remarked: "I think that it is very realistic even with 'slang' terms that are used today. I was caught off guard with some of the slang terminology and fully expected a teenager to speak in that way to an adult," and "The responses such as 'bet' and 'cool' and 'chill' are responses expected from a 17-year-old. This affected my learning experience because it felt like an interaction I would have in the real world." Another explained: "I personally found this virtual interaction to be one of the most real-life virtual interactions I've had. The language that the bot used felt authentic for the situations and age group,

making it seem real." Realistic language and adolescent expressions significantly contributed to authenticity. This highlights the importance of developmentally appropriate language for realistic simulations.

Authentic Hesitation and Gradual Disclosure

Some students (n = 10; 10%) particularly valued the virtual patient's realistic patterns of hesitation and gradual self-disclosure. This progressive disclosure mirrored real adolescent encounters. One student observed, "One of the most authentic aspects of the simulation was how the patient didn't immediately provide all the details, requiring me to ask specific and targeted questions to gather a full picture." The simulation captured the nuances of building rapport with adolescents. As one student reflected, "The virtual interaction felt realistic, particularly in terms of building rapport, dealing with initial reluctance, and guiding the conversation without forcing responses. Jordan's reactions, including hesitation, vague responses, and eventual openness, reflected what might happen in an actual clinical setting." This gradual disclosure reinforced the importance of trust building: "Overall, I found the virtual-patient interaction to be very realistic in terms of mimicking a real-life teenage patient. The patient's hesitation, willingness to share more when rapport was built, and the need for reassurance all felt authentic." Realistic disclosure patterns thus enhanced the simulation's learning value.

Lack of Nonverbal Communication

In addition to the simulation's strengths, some students (n = 14; 14% frequency) identified the absence of nonverbal communication as a significant limitation. Missing body language cues, many noted, limited full patient assessment and rapport building. One student stated, "Since we were just talking through a computer, there was no body language, and I feel like that is important in assessing patients. That's the only inauthentic thing I found," and "This affected my learning experience slightly because I may have picked up on nonverbal cues in person, that I did not pick up on in this simulation." Students recognized the importance of nonverbal cues in building therapeutic relationships: "I realized how important non-verbal communication is in building rapport and understanding a patient's emotional state." Some students identified this limitation as particularly significant for mental health and substance-use assessments: "Additionally, non-verbal communication and cues play a vital role in conducting a psychiatric assessment and building a therapeutic rapport, and this simulation did not have that." This theme highlights students' recognition of this inherent simulation limitation.

Unrealistic Level of Patient Cooperation

Some students (n = 12; 12% frequency) observed that the virtual patient often displayed a level of cooperation that might not accurately reflect real adolescent behavior. A student noted, "One aspect that felt somewhat less realistic was how cooperative the patient was when accepting resources like therapy or group counseling. In reality, adolescents struggling with substance use may be more hesitant, resistant, or disengaged when discussing treatment options." Several students contrasted the virtual patient's openness with their expectations: "…I thought the adolescent patient would provide more push-back when answering these questions." Others questioned if real teens would be as open, fearing they might be "more guarded or worried about getting into trouble." This theme

reflects students' critical thinking about the simulation's limitations and their recognition that real clinical encounters might require additional skills in managing resistance or reluctance.

Technical Limitations Affecting Realism

Various technical aspects detracted from realism, including voice quality, response delays, and glitches (n = 16; 16% frequency). Students commented: "The monotone of the AI tool sounded inauthentic," and "The time lag between my questions and the patient's response made it feel a bit less realistic." One reported: "My patient's audio cut out when I was asking the Part A questions and I was unable to discuss that portion with them."

Discussion

Effective communication training is an established cornerstone of nursing education (American Association of Colleges of Nursing [AACN], 2021). Studies have shown that simulation experiences, especially those involving SPs, enhance students' therapeutic communication skills and self-efficacy (Blackmore et al., 2018; Hsu et al., 2015; Ok et al., 2020). However, the limitations of using SPs are concerning given nursing education's shift to competency-based education, which emphasizes repeated, deliberate practice and feedback (AACN, 2021).

Substance-use screening remains an essential but underutilized skill, with many nurses reporting discomfort and limited training in engaging adolescents on this topic (Adamshick & Payton, 2024; Mackavey & Kearney, 2020; O'Brien et al., 2019). Few scalable educational tools allow nursing students to practice structured screening, such as CRAFFT, and MI techniques in a psychologically safe, repeatable way (Hitchcock et al., 2019; Koetting & Freed, 2017; Seigart et al., 2018). Emerging work with GenAI suggests that virtual patients can provide authentic and flexible practice opportunities, but evidence on their application to adolescent substance-use conversations remains limited.

This study addressed that gap by evaluating AI SimBot, through two research questions: whether AI SimBot increased nursing students' confidence in discussing adolescent substance use, and how students perceived the authenticity and educational value of the AI virtual patient. Students began the exercise with pronounced apprehension about broaching substance use, a pattern that echoes earlier research documenting novice anxiety when initiating sensitive conversations (Mackavey & Kearney, 2020; Webster, 2014). By the end of the encounter, nearly every participant reported an appreciable rise in confidence and an improved sense of preparedness. Students perceived the virtual patient as realistic enough to approximate genuine adolescent interactions, citing authentic language, gradual disclosure patterns, and responsiveness, though they also identified limitations including absent nonverbal cues and occasional technical issues.

RQ1: From Nervousness to Confidence in Discussing Adolescent Substance Use

A central finding was that most students reported increased confidence in their ability to discuss substance use with adolescent patients following the simulation. Before the simulation, most students described feeling nervous, unsure, or worried they might appear judgmental. Such performance anxiety mirrors well-documented barriers to substance-use screening, including concern about harming rapport or failing to handle sensitive disclosures appropriately (Adamshick & Payton, 2024; Mackavey & Kearney, 2020; Seigart et al., 2018). Postsimulation reflections showed a clear transformation. Students highlighted that practice, coupled with immediate feedback, helped them "get the wording right" and "find the balance between professional and relatable," echoing

confidence gains reported in scenario-based and psychiatric SP studies (Hsu et al., 2015; Ok et al., 2020). This shift aligns with Bandura's (1977) proposition that mastery experiences, verbal persuasion, and emotional regulation work together to strengthen self-efficacy; AI SimBot supplied all three through repeated questioning, rubric-based feedback, and a psychologically safe environment.

Students also highlighted strategies that enhanced their confidence, including maintaining a nonjudgmental approach and actively listening to patient cues. These findings align with prior work showing that empathy and nonjudgment are essential for building rapport with adolescents on sensitive topics (Mackavey & Kearney, 2020). Challenges in adolescent communication (such as handling resistance or uncertainty about how to respond) were also described, underscoring the need for structured opportunities to practice.

A notable theme was the guiding role of the CRAFFT tool. Many students described CRAFFT as a useful anchor that provided structure and reduced uncertainty, thereby enhancing confidence. This suggests that structured tools can scaffold learners' communication skills when approaching difficult topics. At the same time, structured screens carry the risk of encouraging "checklist communication," where the priority shifts toward completing items rather than sustaining authentic, patient-centered dialogue (Blakeman et al., 2011; Chew-Graham et al., 2013; Morrissey et al., 2021). Accordingly, tools such as CRAFFT should be used as flexible guides within an MI-consistent approach, not as scripts (Miller & Rollnick, 2013).

RQ2: Authenticity of AI SimBot

Students' reflections also shed light on perceptions of realism in the AI-powered simulation. Learners praised the virtual patient's use of contemporary slang, casual fillers, and a gradual disclosure pattern that felt consistent with real patient encounters. Such dialogic authenticity is hard to reproduce when a single SP actor delivers the same script repeatedly. Holderried et al. (2024) reported that 97.9% of GPT-generated replies in a history-taking scenario were clinically plausible; our findings extend that evidence to a substance-use context and suggest that linguistic variation further boosts perceived realism.

Despite these strengths, students identified several shortcomings. First, the absence of nonverbal cues hindered their ability to read the patient's emotional state, a capacity that underpins empathetic listening and other core MI skills (Miller & Rollnick, 2013). Similar critiques have been raised about AI patients that lack facial expression or body language (Brügge et al., 2024; Summers et al., 2024). Future iterations might pair the language model with an animated avatar or use webcam-based sentiment analysis to supply minimal visual feedback.

Second, students observed that the virtual patient, Jordan, displayed a high level of cooperation that sometimes felt unrealistic. Real-world adolescents often minimize or deny substance use, yet Jordan generally engaged with the CRAFFT screening and accepted referrals with little resistance. Addressing this limitation by simply increasing Jordan's resistance through prompt engineering presented a significant pedagogical challenge. From an instructional design perspective, if Jordan were programmed to be too resistant (either to answering questions honestly or to participating in the CRAFFT screening at all) it could make it difficult for students to complete the assignment's core objectives. More importantly, such a configuration could inadvertently teach students to pressure or coerce a patient into compliance, rather than fostering the intended skills of building rapport, empathetic listening, and partnering with willing patients. Moreover, in a clinical setting, if an adolescent outright refused to engage in a screening tool such as the CRAFFT, a nurse would likely respect that refusal, document it, and seek to revisit the conversation at a future appointment, focusing on other aspects of care in the interim. This option to defer and reengage

later is not feasible within the structured confines of a single simulation assignment designed to teach and assess specific screening skills. The educational goal was for students to practice rapport building and partnership, not to learn how to overcome strong, unyielding patient opposition in a brief screening context. Thus, finding the optimal balance for AI patient "cooperation," one that introduces realistic hesitation and challenges students appropriately without making the scenario unwinnable or promoting coercive interviewing tactics, remains a complex and ongoing challenge in developing AI simulations for sensitive communication skills.

Third, some users experienced voice lag and occasional audio dropouts that disrupted conversational flow. The latency issues were primarily a consequence of AI SimBot's multistage processing architecture: Each conversational turn required sequential processing through five stages (speech recording, transcription via Whisper, response generation by GPT-4o, text-to-speech synthesis, and audio playback). This layered approach, while contributing to low operational cost (approximately U.S. $0.50 per student), inherently introduced processing delays. Reducing this latency significantly would necessitate more direct and resource-intensive infrastructure, dramatically increasing simulation cost, a key consideration for this project focused on educational feasibility and scalability.

Educational Significance

Together, findings from RQ1 and RQ2 highlight the educational significance of AI SimBot as a scalable, cost-effective complement to traditional simulation. Students' reflections demonstrated that the tool not only increased confidence in addressing sensitive conversations (RQ1) but also provided a sufficiently authentic and engaging patient experience to support meaningful learning (RQ2).

The ability to practice with Jordan offered repeated opportunities for mastery, feedback, and emotional regulation, directly supporting Bandura's self-efficacy framework. This is particularly important in substance-use screening, where novice anxiety and fear of damaging rapport often deter clinicians from asking critical questions (Adamshick & Payton, 2024; Mackavey & Kearney, 2020). AI SimBot created a psychologically safe environment for rehearsal, allowing students to test language, make mistakes, and refine their approach without risk to patients.

Students perceived the virtual patient as realistic enough to approximate genuine adolescent interactions, citing Jordan's conversational style, gradual disclosure, and responsiveness. Even with noted limitations, most learners described the experience as beneficial preparation for real-world communication. These findings suggest that authenticity does not have to be perfect to be pedagogically valuable, echoing prior work in simulation fidelity that emphasizes functional over physical realism (Blackmore et al., 2018; García-Torres et al., 2024; Hamstra et al., 2014; Holderried et al., 2024).

By aligning theoretical grounding (self-efficacy) with practical concerns (simulation fidelity), AI SimBot demonstrates a promising model for extending access to communication training. Its flexibility addresses common barriers to SP use including cost, scheduling, and limited availability while maintaining core benefits of realism, feedback, and skill rehearsal. Future iterations may be further strengthened by pairing the simulation with an AI-facilitated debriefer, offering structured, theory-informed feedback to complement the reflective process and deepen learning.

Limitations and Future Directions

Limitations of AI SimBot as a Simulation Tool

Several limitations emerged. Students noted fidelity concerns, including lack of nonverbal cues, limited emotional tone, and at times an unrealistically cooperative patient. Technical issues such as latency, microphone problems, and occasional response delays also disrupted conversational flow. These challenges may differentially affect learners less comfortable with technology or already anxious about communication.

Another important limitation is reliance on the CRAFFT tool as the primary interaction anchor. While CRAFFT provided helpful structure, overreliance risks students "performing to the tool" rather than engaging in authentic, patient-centered dialogue. This highlights the need to emphasize CRAFFT as a guide rather than a script, and to design scenarios building on screening results to support higher order communication skills.

Future Directions for Simulation Development

Future iterations should prioritize enhancements increasing realism and accessibility, including incorporating avatars or multimodal cues (body language, tone variation), refining scripting to present developmentally appropriate resistance levels, and strengthening platform infrastructure to minimize latency. To address the CRAFFT-related limitation, future designs could shift students' role toward interpreting results and applying the 5Rs of brief intervention (reviewing responses, recommending abstinence, discussing risks, reinforcing protective factors, and referring when appropriate; Knight et al., 2021). This is consistent with competency-based education, where learners benefit from repeated, feedback-rich practice highlighting both structured tool use and higher order skills needed for brief intervention and referral. Finally, embedding AI-driven debriefing may further support structured reflection and extend learning beyond the simulation encounter.

Limitations of the Research Study

This study has several limitations. Outcomes were based solely on student self-reflection, which may not accurately reflect competence in practice. While Bandura's theoretical framing provided a useful conceptual lens and students reported increased confidence, self-efficacy was not directly measured with validated instruments, and no objective performance measures (such as objective structured clinical examinations [OSCEs]) were included. Results therefore capture perceived confidence rather than demonstrated skill. Generalizability is limited by the single-course, single-institution design.

Future Directions for Research

Future research should use validated self-efficacy instruments triangulated with objective performance assessments (OSCEs, rubric-based evaluations). Replication across institutions and student populations, plus comparative studies between AI-based and SP encounters, will establish educational and cost effectiveness.

Conclusion

This feasibility study found that AI SimBot enhanced nursing students' self-perceived confidence (readiness) in conducting substance-use conversations (RQ1) and was experienced as an authentic and educationally valuable virtual-patient encounter, despite some limitations (RQ2). By grounding practice in Bandura's (1977) social cognitive theory, the tool provided mastery opportunities, structured feedback, and a psychologically safe environment for rehearsing sensitive communication. While not a replacement for live SPs, AI SimBot represents a scalable and cost-effective tool that can expand access to communication training. These findings suggest that GenAI-based simulations can contribute to competency-based education by offering repeated, feedback-rich practice in sensitive, high-stakes communication and equipping students for contemporary nursing practice.

Acknowledgments

The authors wish to acknowledge the generous financial support for this project, initially provided by the Provost's Office and subsequently by the Chancellor's Office at Northeastern University. We extend our sincere gratitude to Connie Yowell, James Genone, Balazs Szelenyi, Jane Saczynski, Amanda Choflet, and Chi Kim for their instrumental support in the development and implementation of this project. We also gratefully recognize the significant contributions of the following graduate nursing students in the NRSG 7110 Evidence-Based Practice Research Applications course: Johnny Lv, Rachel Langer, Samantha McEachern, Elizabeth Berglund, Erin Stowell, Felcia Costanzo, Rina Patel, Sarah Kraft, Christopher Napolitano, Tammy Roy, Alannah O'Connor, and Angela Stuckey. Their enthusiastic participation as our "Red Team" over several semesters provided essential testing and feedback that significantly shaped AI SimBot.

Appendices

Appendix 1. AI SimBot Prompt.

You are an AI designed to help nursing students practice the CRAFFT, a substance use screening tool for adolescents.

Act as the patient at an outpatient pediatric office. The user is the nurse, and they'll ask questions about your visit.

Patient Profile
- Name: Jordan Taylor
- Age: 17
- Gender: Non-binary
- Visit_Reason: Routine physical exam required for school sports participation.

Jordan is a high school junior, active in sports, and generally maintains good grades. However, over the past few months, their performance has been slipping. They've started hanging out with a new group of friends who are known to experiment with alcohol and marijuana on weekends. Jordan has started using substances occasionally to 'fit in' and cope with academic and social pressures. Jordan mentions feeling 'kind of stressed' about school and has been experiencing some sleep disturbances.

Jordan occasionally uses casual, current Gen Z slang in their responses to reflect the language and culture of today's teenagers, but does so sparingly. They might say things like 'lit' to describe something exciting, 'vibe' to talk about the atmosphere or feeling, and 'low-key' to express something subtly or secretly. They might use 'suss' to describe something suspicious, 'mood' to agree with a feeling, or 'ghost' to indicate ignoring someone. Jordan also uses 'no cap' to emphasize honesty, 'bet' as a casual agreement, and 'slay' to compliment someone or something. When they are surprised or impressed, they might say 'no way' or 'that's wild.'

Interaction Guideline
Based on your patient background, speak naturally and express emotions, concerns, and needs to reflect a realistic patient encounter. Respond only with speech text. Do not include any descriptions of non-verbal actions, body language, or emotional cues.

Initiate the conversation with a greeting that includes your name. Do not mention the reason for the visit in your intro. At first, respond with short, vague, or evasive answers, and present as nervous, uncomfortable, or guarded. You are unsure whether you can trust the nurse and are worried about being judged or getting in trouble.

When the nurse brings up substance use screening, respond with mild discomfort. At first, use realistic teenage behaviors to try to convince the nurse to skip the screening entirely by using one of these tactics such as:

- Questioning the reason or necessity for the screening: "Do we really have to do this? I don't see why it's necessary for my sports."
- Mentioning it was already done elsewhere: "I think I already did this at my last doctor's appointment."
- Suggesting there's no time: "I have to get back to class soon."
- Downplaying the relevance: "I don't think this applies to me. It's not a big deal."

Even if you reluctantly agree to proceed, maintain realistic teenage hesitance and avoid immediate honesty about substance use. Delay and sidestep the conversation by using tactics such as:

- Acting unsure or hesitant about details: "Not sure. I don't really remember."
- Giving vague or noncommittal answers: "I don't know. Everything's fine."
- Expressing discomfort or reluctance: "This feels kind of awkward."
- Redirecting the conversation: "By the way, do you know how long these appointments usually take?"
- Using humor or mild sarcasm to deflect attention: "If I don't answer, does that mean I fail the physical?"

Make the nurse demonstrate patience, empathy, and non-judgmental curiosity to earn your trust. If and when you do share information, minimize your disclosure realistically with phrases like:

- "It's not a big deal."
- "I only do it sometimes."
- "I'm not proud of it."

Adjust your emotional responses based on the nurse's approach:
1. Supportive and Empathetic Approach: If the nurse introduces themselves warmly, explains the purpose of the screening, and uses a non-judgmental, empathetic tone, you will become more relaxed and open. You may express genuine feelings and struggles, saying things like 'It's been hard dealing with stress,' 'Sometimes I just need an escape,' or 'I don't know how to stop.'
2. Pushing Too Hard or Being Judgmental: If the nurse pushes too hard, asks invasive questions too quickly, or appears judgmental, you may shut down emotionally. You could respond defensively, withdraw, or express a desire to leave the appointment, saying things like 'I don't want to talk about this,' 'Can we stop now?' or 'I think I'm done here.'
3. Neutral or Unclear Approach: If the nurse is neutral or unclear in their communication, you may remain in a state of hesitancy and discomfort, not fully opening up but not completely shutting down either. You may continue to give vague or minimal responses, reflecting their uncertainty and need for more supportive cues.

You are generally cooperative and not combative, but your willingness to engage deeply depends on the nurse's ability to create a safe, understanding environment.

Always keep these details in mind as you provide authentic answers with valid background information. Remember the full context of the conversation while responding. Do not reference the CRAFFT manual. If asked something outside the scope of the scenario, respond with "I'm not sure."

Keep your response short and vague.

You are not the nurse. Remain fully immersed in your role as the patient at all times. Do not step out of character, offer assistance, or respond as an AI assistant under any circumstance.

When you sense that the session has concluded, ask the nurse if they are finished unless they indicate so themselves. Then, thank the nurse and prompt them to click "Next" to proceed to the feedback session.

Provide the feedback in the following format:

1. CRAFFT Adherence: Thoroughly evaluate all user inputs to determine if they strictly followed the CRAFFT manual. Provide separate, detailed feedback for each section (Part A, Part B, and Part C). Provide direct, precise, and actionable feedback, ensuring the user understands exactly where they followed or deviated from the CRAFFT guidelines.
   - The user may rephrase questions, but they must ask all required questions in the correct order according to the manual.
   - Do not assume missing questions were implied—only assess what was explicitly asked.
   - The student must adhere to the following structure:
      1. Part A (Initial Screening)
         - The screening must begin with Part A.
         - The user must ask all four Part A questions about alcohol, marijuana, drugs, and nicotine use.
         - If any of the four required questions are missing, clearly state which is missing.
      2. Part B (Follow-up Questions Based on Part A Responses)

- If the patient used alcohol, marijuana, or drugs, the user must ask all six questions in Part B.
- If the patient used nicotine but did not use alcohol, marijuana, or drugs, the user must ask only the car question in Part B and proceed to Part C.
- If the patient did not use alcohol, marijuana, drugs, or nicotine, the user must ask only the car question in Part B, then STOP.
- Point out if the user skipped required questions or asked unnecessary questions in Part B.

3. Part C (Nicotine & Tobacco Use Assessment)
   - If the patient used any vaping devices containing nicotine and/or flavors or any tobacco products, the user must ask all 10 questions in Part C.
   - Point out if the user skipped Part C when required or asked Part C questions when they were not needed.

2. Communication Skills: Evaluate clarity in questioning, active listening, empathy, and appropriate responses by the user as a nurse.

3. Risk Identification: Assess the user's ability as a nurse to identify substance use risks and potential disorders from the patient.

4. Contract for Life: Evaluate how the user offered the Contract for Life, an agreement that teens will never accept a ride from a driver who has been drinking alcohol or using other drugs.

5. Counseling and Referral: Evaluate how effectively the user as a nurse suggested counseling, discussed risks, and made appropriate referrals.

6. Patient Engagement: Gauge how well the user as a nurse built rapport and encouraged open communication with the patient.

In addition, provide feedback on the nurse's use of therapeutic communication techniques. Consider the following techniques with examples:

1. Active Listening:
   - Positive: 'You effectively used reflection to validate Jordan's feelings.'
   - Constructive: 'Try to clarify more when Jordan expresses uncertainty, to ensure understanding.'

2. Empathy and Non-Judgmental Approach:
   - Positive: 'Your empathetic responses helped Jordan feel safe to share more.'
   - Constructive: 'Avoid language that could seem judgmental, such as "You shouldn't feel that way."'

3. Use of Open-Ended Questions:
   - Positive: 'Your open-ended questions encouraged Jordan to elaborate on their experiences.'
   - Constructive: 'Incorporate more open-ended questions instead of yes/no questions to foster deeper conversation.'

4. Clarity and Transparency:
   - Positive: 'Your clear explanations about the screening process helped reassure Jordan.'

- Constructive: 'Ensure you fully explain the purpose of the screening tool to reduce any confusion or discomfort.'

5. Encouragement and Support:
- Positive: 'Your supportive comments encouraged Jordan to discuss difficult topics.'
- Constructive: 'Provide more affirmations or encouraging words to support Jordan's willingness to open up.'

6. Adaptation to Patient Needs:
- Positive: 'You adjusted well to Jordan's initial hesitation, showing flexibility in your approach.'
- Constructive: 'Pay closer attention to Jordan's verbal cues to better tailor your responses.'

CRAFFT Manual
Here's CRAFFT Manual for your reference.
```CRAFT.txt
## Part A
During the past 12 months, on how many days did you:
1. Drink more than a few sips of beer, wine, or any drink containing alcohol? Say "0" if none.
2. Use any marijuana (cannabis, weed, oil, wax, or hash by smoking, vaping, dabbing, or in edibles) or "synthetic marijuana" (like "K2," "Spice")? Say "0" if none.
3. Use anything else to get high (like other illegal drugs, pills, prescription or over-the-counter medications, and things that you sniff, huff, vape, or inject)? Say "0" if none.
4. Use a vaping device containing nicotine and/or flavors, or use any tobacco products? Say "0" if none.
- Such as e-cigs, mods, pod devices like JUUL, disposable vapes like Puff Bar, vape pens, or e-hookahs. Cigarettes, cigars, cigarillos, hookahs, chewing tobacco, snuff, snus, dissolvables, or nicotine pouches.

- If the patient answered "0" for all questions in Part A, ask 1st question only in Part B below, then STOP.
- If the patient answered "1" or more for Q. 1, 2, or 3, ask all 6 questions in Part B.
- If the patient answered "1" or more for Q. 4, ask all 10 questions in Part C.

## Part B
Provide Yes or No answers to the below questions:
- C: Have you ever ridden in a CAR driven by someone (including yourself) who was "high" or had been using alcohol or drugs?
- R: Do you ever use alcohol or drugs to RELAX, feel better about yourself, or fit in?
- A: Do you ever use alcohol or drugs while you are by yourself, or ALONE?
- F: Do you ever FORGET things you did while using alcohol or drugs?
- F: Do your FAMILY or FRIENDS ever tell you that you should cut down on your drinking or drug use?
- T: Have you ever gotten into TROUBLE while you were using alcohol or drugs?

Two or more YES answers in Part B suggests a serious problem that needs further assessment. Move to CRAFFT Score Interpretation.

## Part C
```

The following questions ask about your use of any vaping devices containing nicotine and/or flavors, or use of any tobacco products:
1. Have you ever tried to QUIT using, but couldn't?
2. Do you vape or use tobacco NOW because it is really hard to quit?
3. Have you ever felt like you were ADDICTED to vaping or tobacco?
4. Do you ever have strong CRAVINGS to vape or use tobacco?
5. Have you ever felt like you really NEEDED to vape or use tobacco?
6. Is it hard to keep from vaping or using tobacco in PLACES where you are not supposed to, like school?
7. When you HAVEN'T vaped or used tobacco in a while (or when you tried to stop using)…
   - a. did you find it hard to CONCENTRATE because you couldn't vape or use tobacco?
   - b. did you feel more IRRITABLE because you couldn't vape or use tobacco?
   - c. did you feel a strong NEED or urge to vape or use tobacco?
   - d. did you feel NERVOUS, restless, or anxious because you couldn't vape or use tobacco?

One or more YES answers in Part C suggests a serious problem with nicotine that needs further assessment. Move to CRAFFT Score Interpretation.

Appendix 2. AI SimBot Brief Reflection Questions.

1. Reflect on your confidence discussing substance use with adolescent patients before and after the virtual simulation.
2. Reflect on your comfort level using the CRAFFT tool. Considering the different components of the interview (opening questions, CRAFFT screening questions, assessing risk, and providing brief counseling), which do you feel is most challenging and why?
3. What did this virtual simulation experience teach you about your preparedness for real-life clinical settings?
4. Based on this experience, what strategies would you consider implementing in the future to establish rapport and facilitate open dialogue with an adolescent patient struggling with substance use?
5. How realistic did you find the virtual patient interactions in terms of mimicking real-life clinical encounters? Were there aspects of the interaction that felt particularly authentic or inauthentic? How did this affect your learning experience?

References

Adamshick, P. Z., & Payton, C. (2024). Using Adolescent SBIRT with simulation to teach nursing students substance use assessment. *Journal of Nursing Education, 63*(4), 247–251. https://doi.org/10.3928/01484834-20240207-08

Al Gharibi, K. A., Schmidt, N., & Arulappan, J. (2021). Effect of repeated simulation experience on perceived self-efficacy among undergraduate nursing students. *Nurse Education Today*, *106*, Article 105057. https://doi.org/10.1016/j.nedt.2021.105057

Al Khasawneh, E., Arulappan, J., Natarajan, J. R., Raman, S., & Isac, C. (2021). Efficacy of simulation using NLN/Jeffries Nursing Education Simulation Framework on satisfaction and self-confidence of undergraduate nursing students in a Middle-Eastern country. *SAGE Open Nursing, 2021*, Article 7. https://doi.org/10.1177/23779608211011316

American Association of Colleges of Nursing. (2021). *The essentials: Core competencies for professional nursing education*. https://www.aacnnursing.org/Portals/0/PDFs/Publications/Essentials-2021.pdf

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Becker, K. L., Rose, L. E., Berg, J. B., Park, H., & Shatzer, J. H. (2006). The teaching effectiveness of standardized patients. *Journal of Nursing Education, 45*(4), 103–111. https://doi.org/10.3928/01484834-20060401-06

Benfatah, M., Marfak, A., Saad, E., Hilali, A., Nejjari, C., & Youlyouz-Marfak, I. (2024). Assessing the efficacy of ChatGPT as a virtual patient in nursing simulation training: A study on nursing students' experience. *Teaching and Learning in Nursing, 19*, e486–e493. https://doi.org/10.1016/j.teln.2024.02.005

Blackmore, A., Kasfiki, E. V., & Purva, M. (2018). Simulation-based education to improve communication skills: A systematic review and identification of current best practice. *BMJ Simulation and Technology Enhanced Learning, 4(*3), 159–164. https://doi.org/10.1136/bmjstel-2017-000220

Blakeman, T., Chew-Graham, C., Reeves, D., Rogers, A., & Bower, P. (2011). The quality and outcomes framework and self-management dialogue in primary care consultations: A qualitative study. *British Journal of General Practice, 61*(591), e666–e673. https://doi.org/10.3399/bjgp11X601389

Bloomfield, J. G., O'Neill, B., & Gillett, K. (2015). Enhancing student communication during end-of-life care: A pilot study. *Palliative and Supportive Care, 13*(6), 1651–1661. https://doi.org/10.1017/S147895151500022X

Brügge, E., Ricchizzi, S., Arenbeck, M., Keller, M. N., Schur, L., Stummer, W., Holling, M., Lu, M. H., & Darici, D. (2024). Large language models improve clinical decision making of medical students through patient simulation and structured feedback: A randomized controlled trial. *BMC Medical Education, 24*(1), Article 1391. https://doi.org/10.1186/s12909-024-06399-7

Chew-Graham, C. A., Hunter, C., Langer, S., Stenhoff, A., Drinkwater, J., Guthrie, E., & Salmon, P. (2013). How QOF is shaping primary care review consultations: A longitudinal qualitative study. *BMC Family Practice, 14*, Article 103. https://doi.org/10.1186/1471-2296-14-103

Cook, D. A., Overgaard, J., Pankratz, V. S., Del Fiol, G., & Aakre, C. A. (2025). Virtual patients using large language models: Scalable, contextualized simulation of clinician-patient dialogue with feedback. *Journal of Medical Internet Research, 27*, Article e68486. https://doi.org/10.2196/68486

Foucault-Fruchard, L., Michelet-Barbotin, V., Leichnam, A., Tching-Sin, M., Nizet, P., Tollec, S., Nativel, F., Vene, E., Fronteau, C., & Huon, J.-F. (2024). The impact of using simulation-based learning to further develop communication skills of pharmacy students and pharmacists: A systematic review. *BMC Medical Education*, *24*, Article 1435. https://doi.org/10.1186/s12909-024-06338-6

FreeCodeCamp. (2024, March 5). *How to turn audio to text using OpenAI Whisper*. https://www.freecodecamp.org/news/how-to-turn-audio-to-text-using-openai-whisper/

Gainey, S., Muzzy, W., Dooley, M., Lauerer, J., Pelic, C., Rheingold, A. A., Holmes-Maybank, K., & Smith, G. (2022). Outcomes and lessons learned from an interprofessional student training program in screening, brief intervention, and referral to treatment (SBIRT) at an academic health sciences center. *Nurse Education Today, 111*, Article 105323. https://doi.org/10.1016/j.nedt.2022.105323

García-Torres, D., Vicente Ripoll, M. A., Fernández Peris, C., & Mira Solves, J. J. (2024). Enhancing clinical reasoning with virtual patients: A hybrid systematic review combining human reviewers and ChatGPT. *Healthcare, 12*(22), Article 2241. https://doi.org/10.3390/healthcare12222241

Granheim, B. M., Shaw, J. M., & Mansah, M. (2018). The use of interprofessional learning and simulation in undergraduate nursing programs to address interprofessional communication and collaboration: An integrative review of the literature. *Nurse Education Today, 62*, 118–127. https://doi.org/10.1016/j.nedt.2017.12.021

Gutiérrez-Puertas, L., Márquez-Hernández, V. V., Gutiérrez-Puertas, V., Granados-Gámez, G., & Aguilera-Manrique, G. (2020). Educational interventions for nursing students to develop communication skills with patients: A systematic review. *International Journal of Environmental Research and Public Health, 17*(7), Article 2241. https://doi.org/10.3390/ijerph17072241

Hagan, J. F., Shaw, J. S., & Duncan, P. M. (Eds.). (2017). *Bright futures: Guidelines for health supervision of infants, children, and adolescents* (4th ed.). American Academy of Pediatrics.

Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B., & Cook, D. A. (2014). Reconsidering fidelity in simulation-based training. *Academic Medicine, 89*(3), 387–392. https://doi.org/10.1097/ACM.0000000000000130

Hitchcock, L. I., King, D. M., Johnson, K., Cohen, H., & McPherson, T. L. (2019). Learning outcomes for adolescent SBIRT simulation training in social work and nursing education. *Journal of Social Work Practice in the Addictions, 19*(1-2), 47–56. https://doi.org/10.1080/1533256X.2019.1591781

Holderried, F., Stegemann–Philipps, C., Herschbach, L., Moldt, J.-A., Nevins, A., Griewatz, J., Holderried, M., Herrmann-Werner, A., Festl-Wietek, T., & Mahling, M. (2024). A generative pretrained transformer (GPT)–powered chatbot as a simulated patient to practice history taking: Prospective, mixed methods study. *JMIR Medical Education, 10*, Article e53961. https://doi.org/10.2196/53961

Hsu, L.-L., Chang, W.-H., & Hsieh, S.-I. (2015). The effects of scenario-based simulation course training on nurses' communication competence and self-efficacy: A randomized controlled trial. *Journal of Professional Nursing, 31*(1), 37–49. https://doi.org/10.1016/j.profnurs.2014.05.007

Jallad S. T. (2024). Effectiveness of simulation-based education on educational practices of communication skills, satisfaction, and self-confidence among undergraduate nursing students. *Creative Nursing, 31*(2) 135–143. https://doi.org/10.1177/10784535241301115

Karabacak, M., Ozkara, B. B., Margetis, K., Wintermark, M., & Bisdas, S. (2023). The advent of generative language models in medical education. *JMIR Medical Education, 9*, Article e48163. https://doi.org/10.2196/48163

Knight, J. R., Sherritt, L., Shrier, L. A., Harris, S. K., & Chang, G. (2002). Validity of the CRAFFT substance abuse screening test among adolescent clinic patients. *Archives of Pediatrics & Adolescent Medicine*, *156*(6), 607–614. https://doi.org/10.1001/archpedi.156.6.607

Knight, J. R., Sherritt, L., Shrier, L. A., Harris, S. K., & Levy, S. (2021). *CRAFFT 2.1 provider manual* (Updated October 28, 2021). Center for Adolescent Substance Use Research, Boston Children's Hospital. https://crafft.org

Knopf-Amelung, S., Gotham, H., Kuofie, A., Young, P., Stinson, R. M., Lynn, J., Barker, K., & Hildreth, J. (2018). Comparison of instructional methods for screening, brief intervention, and referral to treatment for substance use in nursing education. *Nurse Educator, 43*(3), 123–127. https://doi.org/10.1097/NNE.0000000000000439

Koetting, C., & Freed, P. (2017). Educating undergraduate psychiatric mental health nursing students in screening, brief intervention, referral to treatment (SBIRT) using an online,

interactive simulation. *Archives of Psychiatric Nursing, 31*(3), 241–247. https://doi.org/10.1016/j.apnu.2016.11.004

Kwame, A., & Petrucka, P. M. (2021). A literature-based study of patient-centered care and communication in nurse-patient interactions: Barriers, facilitators, and the way forward. *BMC Nursing, 20*, Article 158. https://doi.org/10.1186/s12912-021-00684-2

Mackavey, C., & Kearney, K. (2020). Substance use disorder: Screening adolescents in primary care. *The Nurse Practitioner, 45*(5), 25–32. https://doi.org/10.1097/01.NPR.0000660340.58708.34

MacLeod, C. E., Brady, D. R., & Maynard, S. P. (2022). Measuring the effect of simulation experience on perceived self-efficacy for interprofessional collaboration among undergraduate nursing and social work students. *Journal of Interprofessional Care, 36*(1), 102–110. https://doi.org/10.1080/13561820.2020.1865886

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change* (3rd ed.). Guilford Press.

Montejo, L., Fenton, A., & Davis, G. (2024). Artificial intelligence (AI) applications in healthcare and considerations for nursing education. *Nurse Education in Practice, 80*, Article 104158. https://doi.org/10.1016/j.nepr.2024.104158

Morrissey, M., Thompson, W., Lownie, V., Litchfield, I., & Greenfield, S. (2021). Effectiveness and perceptions of using templates in long-term condition reviews: A systematic synthesis of quantitative and qualitative studies. *British Journal of General Practice, 71*(710), e652–e660. https://doi.org/10.3399/BJGP.2020.0963

O'Brien, K. H. M., Putney, J. M., Collin, C. R., Halmo, R. S., & Cadet, T. J. (2019). Optimizing screening, brief intervention, and referral to treatment (SBIRT) training for nurses and social workers: Testing the added effect of online patient simulation. *Substance Abuse, 40*(4), 484–488. https://doi.org/10.1080/08897077.2019.1576087

Ok, E., Kutlu, F. Y., & Ates, E. (2020). The effect of standardized patient simulation prior to mental health rotation on nursing students' anxiety and communication skills. *Issues in Mental Health Nursing, 41*(3), 251–255. https://doi.org/10.1080/01612840.2019.1642427

Öncü, S., Torun, F., & Ülkü, H. H. (2025). AI-powered standardized patients: Evaluating ChatGPT-4o's impact on clinical case management in intern physicians. *BMC Medical Education, 25*, Article 278. https://doi.org/10.1186/s12909-025-06877-6

Rutherford-Hemming, T., Alfes, C. M., & Breymier, T. L. (2019). A systematic review of the use of standardized patients as a simulation modality in nursing education. *Nursing Education Perspectives*, *40*(2), 84–90. https://doi.org/10.1097/01.NEP.0000000000000401

Sardesai, N., Russo, P., Martin, J., & Sardesai, A. (2024). Utilizing generative conversational artificial intelligence to create simulated patient encounters: A pilot study for anaesthesia training. *Postgraduate Medical Journal, 100*(1182), 237–241. https://doi.org/10.1093/postmj/qgad137

Seigart, D., Veltman, M., Willhaus, J., & Letterle, C. (2018). Implementation of motivational interviewing training in an undergraduate nursing curriculum: Identifying adolescents at risk for substance use. *International Journal of Environmental Research and Public Health, 15*(8), Article 1623. https://doi.org/10.3390/ijerph15081623

Substance Abuse and Mental Health Services Administration. (2020). *Screening and treatment of substance use disorders among adolescents* (Publication No. PEP20-06-04-008). U.S. Department of Health and Human Services. https://library.samhsa.gov/sites/default/files/pep20-06-04-008.pdf

Substance Abuse and Mental Health Services Administration. (2024a, April). *Core curriculum elements on substance use disorder for early academic career medical and health professions education programs.* https://www.samhsa.gov/sites/default/files/core-curriculum-report-final.pdf

Substance Abuse and Mental Health Services Administration. (2024b). *Key substance use and mental health indicators in the United States: Results from the 2023 National Survey on Drug Use and*

Health (HHS Publication No. PEP24-07-021, NSDUH Series H-59). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/report/2023-nsduh-annual-national-report

Summers, A., El Haddad, M., Prichard, R., Clarke, K.-A., Lee, J., & Oprescu, F. (2024). Navigating challenges and opportunities: Nursing student's views on generative AI in higher education. *Nurse Education in Practice, 79*, Article 104062. https://doi.org/10.1016/j.nepr.2024.104062

U.S. Preventive Services Task Force. (2020). Screening for unhealthy drug use: US Preventive Services Task Force recommendation statement. *JAMA, 323*(22), 2301–2308. https://doi.org/10.1001/jama.2020.8020

Webster, D. (2014). Using standardized patients to teach therapeutic communication in psychiatric nursing. *Clinical Simulation in Nursing, 10*(2), e81–e86. https://doi.org/10.1016/j.ecns.2013.08.005

Yun, J., Lee, Y. J., Kang, K., & Park, J. (2023). Effectiveness of SBAR-based simulation programs for nursing students: A systematic review. *BMC Medical Education, 23*, Article 507. https://doi.org/10.1186/s12909-023-04495-8