

Reimagining the SP in the Era of Artificial Intelligence

Andrew B. Nevins

anevins@stanford.edu

Stanford University School of Medicine

Karen Thomson Hall

Stanford University School of Medicine

David M. Gaba

Stanford University School of Medicine

Thomas J. Caruso

Stanford University School of Medicine

advancing-simulation-practice

Keywords: Standardized patient, artificial intelligence, simulation-based education, standardized patient methodology, deliberate practice, health professions education

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11013320/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

For more than five decades, standardized or simulated patients (SPs) have been central to simulation-based clinical skills education. Rapid advances in artificial intelligence (AI) now enable responsive, scalable simulated patient encounters, creating an inflection point for SP methodology. Rather than viewing AI as a disruption to SP practice, we propose a broader conceptualization of SP methodology in which its defining educational principles can be enacted through different modalities with distinct educational strengths. Human SPs remain essential for experiences requiring interpersonal interaction, emotional nuance, physical examination, and many assessment applications. AI SPs offer complementary strengths through structured deliberate practice, individualized feedback, and longitudinal monitoring of learner progress. Thoughtful integration and professional oversight will be essential to preserve the rigor, ethics, and human-centered values of SP practice.

Full Text

For more than half a century, standardized or simulated patients (SPs) have been a cornerstone of health professions education. SPs were introduced to enhance rigor, consistency, and authenticity to clinical skills teaching and assessment (1, 2). Early applications focused on observing learners' ability to obtain medical histories and perform focused physical examinations in a controlled yet realistic environment, addressing concerns about variability and subjectivity in clinical evaluation (3-7).

The role of SPs expanded over subsequent decades. What began as a tool for teaching and formative feedback became central to curriculum design and summative assessment. SP encounters became integral to objective structured clinical examinations (OSCEs) (8-10), enabling educators to assess not only data gathering and examination skills, but also communication, professionalism, clinical reasoning, and clinical decision-making. The educational validity and reliability of SP-based assessments are well established, supporting the widespread adoption of SP programs across medical and other health professions curricula (11-15).

Perhaps the most visible institutional endorsement of SP methodology has been its incorporation into high-stakes licensure and certification assessments. For nearly two decades, the United States Medical Licensing Examination Step 2 Clinical Skills (CS) examination relied on SPs to assess core clinical competencies (16). Although Step 2 CS was discontinued in 2020 (17), comparable approaches continue internationally. For example, the Medical Council of Canada's National Assessment Collaboration (NAC) Examination uses SPs in a national OSCE, and the Pharmacy Examining Board of Canada incorporates simulated participants into its national qualifying examinations (18, 19). Beyond physician and pharmacy licensure, SP-based assessments remain widely used across postgraduate medical education and specialty certification worldwide, underscoring the enduring value of this methodology for assessing communication, clinical reasoning, professionalism, and other essential clinical competencies.

The educational landscape in which SP programs operate is undergoing a significant transformation. Advances in artificial intelligence (AI), particularly in natural language processing and generative models,

now enable highly interactive, responsive, and increasingly realistic patient portrayals (20-22). These systems engage learners in open-ended clinical dialogue, adapt to learner input in real time, and deliver encounters at a scale difficult to achieve using human SPs alone. (23-25) As a result, many institutions are exploring or adopting AI SPs to supplement existing educational activities (26).

Rather than viewing this moment as a disruption that threatens traditional SP programs, it is instead an opportunity to reconsider how SP methodology functions in contemporary health professions education. We propose reconceptualizing SP methodology to include both human and AI modalities as integrated components of contemporary simulation programs, each contributing distinct strengths to clinical skills education while preserving the unique value of human SPs. AI SPs are not positioned to replace traditional SPs but rather extend a long-standing and evolving educational tradition.

Core principles of SPs

Although SPs are traditionally defined as trained individuals who portray patients, the educational construct is fundamentally characterized by the ability to faithfully reproduce an authentic patient encounter in support of education, assessment, and research rather than by the identity of the performer itself (27). Traditional SPs are trained to consistently present a clinical scenario, thereby allowing for meaningful learning, practice, and assessment. The strengths of traditional SPs lie in their human presence: their ability to convey emotion, respond dynamically, and provide nuanced feedback from the patient's perspective in a safe, supportive environment. These features have made SPs particularly valuable in assessing communication skills, empathy, and professionalism.

To further underscore the importance of SPs in medical education, the Association of SP Educators (ASPE) serves as the international professional organization dedicated to advancing SP methodology through shared expertise and best practices (28). Central to its work are the ASPE Standards of Best Practice, a consensus-based framework outlining core values (quality, safety, professionalism, accountability, and collaboration) and the domains of practice – case development, SP training, program management, professional development, and maintenance of a safe learning environment (29, 30). Together, these standards help provide a framework for promoting rigorous, high-quality SP programs and educational experiences.

The terminology used to describe these educational roles has evolved alongside the field itself. Although the term “standardized patient” remains common in the United States, “simulated patient” is more widely used internationally, and contemporary simulation practice increasingly recognizes broader roles such as simulated participants. Consistent with the terminology adopted by ASPE, we use the abbreviation SP throughout this manuscript as an inclusive term encompassing standardized patients, simulated patients, and simulated participants.

Despite their established role and the rigor provided by professional standards, traditional SP programs face practical challenges. They are resource-intensive, requiring ongoing recruitment, training, and quality assurance. Moreover, the degree of standardization expected of SPs should be determined by the

educational purpose rather than viewed as an absolute objective. While high-stakes assessments require carefully calibrated and consistent portrayals, formative teaching encounters often benefit from allowing experienced SPs to exercise professional judgment and respond authentically to learners, provided they remain faithful to the educational objectives of the case. These considerations have prompted educators to explore complementary modalities that preserve core SP principles while addressing scalability and access.

The emergence of AI SPs

Prior to the emergence of AI SPs, many programs experimented with “virtual standardized patients,” computer-based avatars that allowed learners to gather histories and obtain clinical information by typing questions or selecting from predefined menus (31). These platforms expanded access to cases and provided repeatable practice, but interactions were constrained by scripted logic and limited responsiveness, which restricted the realism of learner-patient interactions. Accordingly, these systems were typically used as adjunctive learning tools rather than as substitutes for live SP encounters. Nonetheless, virtual SPs were a crucial developmental step, laying the technological and pedagogical groundwork for today’s AI-driven systems.

AI SPs represent the latest evolution in simulated clinical encounters. Advances in AI now enable dynamic, conversational, and adaptive interactions that more closely approximate responsive patient behavior. Unlike earlier computer-based simulations or branching logic cases (32, 33), contemporary AI SPs can engage in open-ended dialogue, respond flexibly to learner input, and adapt in real time. These systems include text-based conversational agents, voice-enabled avatars, or immersive virtual environments that place learners in three-dimensional clinical scenarios.

A key advantage of AI SPs is their potential to deliver highly consistent, repeatable interactions. When carefully designed, prompted, tested, and monitored, AI SPs have the potential to present the same case to large numbers of learners without fatigue or scheduling constraints. Their digital architecture also offers the possibility of automatically capturing detailed interaction data, creating new opportunities for individualized feedback and monitoring learner progress over time. However, such consistency is not an inherent property of AI systems. Performance may vary substantially across large language models and prompting strategies, making rigorous prompt engineering, iterative validation, and ongoing quality assurance essential if AI SPs are to provide reproducible, standardized simulations.

Educational strengths and limitations of human and AI SPs

Both human and AI SPs provide substantial educational value, grounded in the shared principles of standardization and intentional instructional design. Traditional SPs offer strengths that remain difficult to technologically replicate. Their human presence allows learners to engage with authentic emotion, nuanced nonverbal communication, and the unpredictability of interpersonal interaction – features particularly important for teaching empathy and professionalism. Human SPs can also assess physical examination skills, evaluating whether maneuvers are performed correctly and with appropriate attention

to patient comfort. In addition, they provide rich, qualitative feedback from the patient's perspective, while their realism supports credibility and face validity in summative assessments.

AI SPs, in contrast, excel in their scalability and consistency. Their on-demand availability supports repeated practice and deliberate rehearsal, making them particularly useful for early skills development, history-taking practice, and remediation in low-stakes environments. High levels of standardization allow for consistent learner experiences, while automated feedback and performance metrics support longitudinal tracking.

Despite their considerable promise, AI SPs have important limitations that currently constrain their role in health professions education. One of the principal challenges is achieving sufficient authenticity and reliability to support realistic clinical encounters. Although AI SPs continue to improve rapidly, their performance remains dependent on the underlying large language model, system design, and prompting strategies, and they cannot yet fully replicate the nuanced emotional responses, nonverbal communication, or physical interactions of human SPs. In addition, AI systems remain susceptible to hallucinations, inconsistent responses across repeated interactions, and other forms of generative instability that may compromise educational reliability if not carefully monitored. There is also a risk that learners may become overly reliant on AI-generated feedback or adapt their communication to the expectations of the system rather than to authentic patient interactions. Most notably, AI SPs are currently unable to directly assess learners' hands-on physical examination techniques or other tactile clinical skills. Consequently, AI SPs are particularly well suited for formative assessment and deliberate practice, where repeated exposure, immediate automated feedback, and scalability offer substantial educational advantages. Human SPs, however, remain indispensable for educational activities requiring authentic interpersonal interactions, direct observation of communication and professionalism, and assessment of hands-on physical examination skills, including many summative and high-stakes assessments.

Together, human and AI SPs offer complementary educational value, reinforcing the importance of aligning each modality with specific learning objectives. (Table 1)

Complementary roles in contemporary SP practice

The critical question is not whether human or AI SPs are superior, but which modality, or combination of modalities, is best suited for the intended learning objectives. Although human SPs and AI SPs are distinct educational modalities, both rely on the thoughtful application of SP methodology, including intentional case design, alignment with clearly defined learning objectives, appropriate calibration, and ongoing quality assurance. Each offers complementary strengths that can be matched to different educational needs. Human SPs remain essential where interpersonal nuance, emotional realism, and direct observation of communication or physical examination skills are required, whereas AI SPs can expand opportunities for structured deliberate practice, reinforce foundational skills, and facilitate longitudinal learner monitoring. As AI-enabled patient simulation continues to evolve, important questions remain regarding how established SP standards and governance frameworks should be

adapted to systems that incorporate automated interactions, performance assessment, and feedback. Nevertheless, because both modalities draw upon the core principles of SP methodology, SP educators are uniquely positioned to lead these discussions and guide their design, implementation, quality assurance, and responsible integration of AI SPs across the curriculum.

We propose a broader conceptualization of SP methodology that distinguishes its defining principles from the modality through which they are enacted, recognizing that this represents a proposed expansion of SP methodology rather than an established consensus within simulation-based education. Although human performance and relational interaction remain central to traditional SP practice, we suggest that SP methodology is also grounded in educational intent, deliberate case design, standardization, and alignment with learning objectives. This expanded perspective accommodates both human and AI SPs while preserving the educational value of each.

This broader conceptualization carries important implications for program leadership, faculty development, and scholarship. SP programs, drawing on decades of expertise in case design and assessment, are uniquely positioned to guide the thoughtful integration of human and AI modalities. Collaboration among educators, technologists, and SP professionals will be essential to ensure implementation that is educationally rigorous, ethically sound, and pedagogically aligned.

Concerns that AI SPs might replace human SPs reflect a misunderstanding of their role: AI SPs are tools, not educators. Their effectiveness depends on thoughtful design, alignment with learning objectives, faculty oversight, and ongoing evaluation. Ethical considerations, including bias, transparency, and data privacy, must be addressed, and learners should be clearly informed about the nature and purpose of AI encounters.

As AI SP technologies mature, SP programs may be well served by thoughtfully incorporating these tools into their existing infrastructures rather than viewing them as competing entities. SP programs already possess deep expertise in designing authentic learning experiences, developing cases, providing feedback, and implementing educational assessment. This expertise provides an essential foundation for the thoughtful adaptation of these principles to AI SPs. Thoughtful integration of human and AI SPs preserves the pedagogical rigor and ethical standards that define high-quality simulation.

If SP methodology is to evolve responsibly, AI SPs cannot remain outside established professional standards. As AI-enabled patient simulation evolves, it should not develop as a parallel pathway outside SP methodology. Instead, its responsible integration should be guided by the accumulated expertise, standards, and educational principles developed through decades of SP practice. The same commitment to quality, safety, and accountability that defines human SP practice must now be extended – through formalized guidelines and leadership analogous to ASPE’s role in traditional SP methodology – to AI SP design and deployment. The five domains of the ASPE Standards of Best Practice provide a useful framework for comparing human and AI SP roles and for considering how similar standards might guide their responsible integration into simulation programs (Table 2).

Rather than competing approaches, human and AI SPs should be viewed as complementary modalities with distinct educational strengths. When thoughtfully integrated within SP programs, they have the potential to strengthen clinical skills education while preserving the pedagogical rigor, humanistic values, and educational principles that have long defined SP methodology.

Abbreviations

SP

standardized patient

AI

artificial intelligence

OSCE

objective structured clinical examination

ASPE

Association of SP (standardized patient) Educators

CS

Clinical skills

NAC

National Assessment Collaboration

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and materials: Not applicable

Competing interests: None

Funding: None

Author's contributions: All authors contributed to the development and refinement of the content, drafting and revision of the manuscript, and provided final approval of the manuscript as submitted.

Acknowledgements in the writing of this paper: AI was used to create the Graphical Abstract. Otherwise, no AI tools were used in the preparation of this manuscript.

References

1. Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. J Med Educ 1964;39:802-5. PMID: 14180699.

2. Barrows HS. Simulated patients in medical teaching. *Can Med Assoc J.* 1968;98(14):674–6. PMID: 5646104; PMCID: PMC1924019.
3. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med.* 1993;68(6):443–53.
4. Ainsworth MA, Rogers LP, Markus JF, Dorsey NK, Blackwell TA, Petrusa ER. Standardized patient encounters: a method for teaching and evaluation. *JAMA.* 1991;266(10):1390–6.
5. Stillman PL, Swanson DB, et al. Assessing clinical skills of residents with standardized patients. *Ann Intern Med.* 1986;105(5):762–71.
6. Hardee JT, Kasper IK. From standardized patient to care actor: evolution of a teaching methodology. *Perm J.* 2005 Summer;9(3):79–82. 10.7812/TPP/05-030. . PMID: 22811638; PMCID: PMC3396073.
7. Flanagan OL, Cummings KM. Standardized Patients in Medical Education: A Review of the Literature. *Cureus.* 2023;15(7):e42027. 10.7759/cureus.42027. . PMID: 37593270; PMCID: PMC10431693.
8. Harden RM. What is an OSCE? *Med Teach.* 1988;10:19–22.
9. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE guide 81. Part I: An historical and theoretical perspective. *Med Teach.* 2013;35:e1437–46.
10. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE guide 81. Part II: organisation & administration. *Med Teach.* 2013;35:e1447–63.
11. Mast TA, Barrows HS, editors. Special section: Annex to the proceedings of the AAMC Consensus Conference on the use of standardized patients in the teaching and evaluation of clinical skills. *Teach Learn Med.* 1994;6:1.
12. Miller GE. The assessment of clinical skills/ competence/performance. *Acad Med.* 1990;65(9 suppl):S63–7.
13. Anderson MB, Stillman PL, Wang Y. Growing use of standardized patients in teaching and evaluation in medical education. *Teach Learn Med.* 1994;6:15–22.
14. Barrows HS, Williams RG. A comprehensive performance-based assessment of fourth year students clinical skills. *J Med Ed.* 1987;62:805–9.
15. Shirar LE, Vu NV, Colliver JA, Barrows HS. A survey of study methods, preparation time, test-taking strategies, and perceptions of test validity on a clinical performance- based examination. *Acad Med.* 1992;67(10 suppl):S10–2.
16. Haist SA, Katsufakis PJ, Dillon GF. The evolution of the United States Medical Licensing Examination (USMLE): Enhancing assessment of practice-related competencies. *JAMA.* 2013;310:2245–6.
17. United States Medical Licensing Examination. <https://www.usmle.org/work- relaunch-usmle-step-2-cs-discontinued>. Accessed March 1, 2026.

18. Medical Council of Canada. NAC Examination. <https://mcc.ca/examinations-assessments/nac-examination/>. Accessed July 13, 2026.
19. Pharmacy Examining Board of Canada. https://pebc.ca/pharmacists/qualifying-examination/preparing-for-the-examination/part-i-2/?utm_source=chatgpt.com. Accessed July 13, 2026.
20. Maicher KR, Stiff A, Scholl M et al. Artificial intelligence in virtual standardized patients: Combining natural language understanding and rule-based dialogue management to improve conversational fidelity. *Med Teach*. 2022.
21. Yamamoto A, Koda M, Ogawa H et al. Enhancing Medical Interview Skills Through AI-Simulated Patient Interactions: Nonrandomized Controlled Trial. *JMIR Med Educ*. 2024.
22. Yu H, Zhou J, Li L et al. Simulated patient systems powered by large language model-based AI agents offer potential for transforming medical education. *Commun Med (Lond)*. 2025.
23. García-Torres D, Vicente Ripoll MA, Fernández Peris C, Mira Solves JJ. Enhancing clinical reasoning with virtual patients: A hybrid systematic review combining human reviewers and ChatGPT. *Healthcare (Basel)*; 2024.
24. Jo Y-W, Lee M, Yang H-J. Large language model-based virtual patient simulations in medical and nursing education: A review. *Appl Sci*. 2025.
25. Zidoun Y, El Mardi A. Artificial Intelligence (AI)-Based simulators versus simulated patients in undergraduate programs: A protocol for a randomized controlled trial. *BMC Med Educ*. 2024.
26. Almansour M, Soliman M, Aldekhyyel R et al. An academic viewpoint on the integration of generative artificial intelligence in medical education. *Cureus*. 2025.
27. Lioce L, Lopreiato J, Anderson M, et al. editors. and the Terminology and Concepts Working Group. Rockville, MD: Healthcare Simulation Dictionary, Third Edition. Agency for Healthcare Research and Quality; Society for Simulation in Healthcare, 2025.
28. Association of Standardized Patient Educators (ASPE). *About ASPE*. Association of Standardized Patient Educators website. [website. https://www.aspeducators.org/about-aspe](https://www.aspeducators.org/about-aspe). Accessed March 1, 2026.
29. Lewis K, Bohnert C, Gammon W, et al. The Association of Standardized Patient Educators (ASPE) Standards of Best Practice (SOBP). *Adv Simul*. 2017;2:10.
30. Association of Standardized Patient Educators (ASPE). *Standards of Best Practice (SOBP)*. <https://www.aspeducators.org/standards-of-best-practice-> Accessed March 1, 2026.
31. Maicher K, Danforth D, Price A, et al. Developing a conventional virtual standardized patient to enable students to practice history-taking skills. *Simul Healthc*. 2018;12(2):124–31.
32. Plackett R, Kassianos AP, Mylan S, et al. The effectiveness of using virtual patient educational tools to improve medical students' clinical reasoning skills: a systematic review. *BMC Med Educ*. 2022;22(1):365.

33. Kononowicz AA, Woodham LA, Edelbring S, et al. Virtual Patient Simulations in Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res. 2019;21(7):e14676.

Tables

Tables are available in the Supplementary Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [NEVINSTable1.docx](#)
- [NEVINSTable2.docx](#)
- [GraphicalAbstract1.png](#)