

Use of a Virtual Standardized Patients for Facilitating Conversational Medical Interview **Tugba Tokel, PhD, Thomas B. Talbot, MD, Merrick Bautista, BS, Andrew Leeds²**

Abstract

The purpose of this study was to examine the effects of USC Standard Patient, medical interviewing tool with virtual standardized patients (VSP). Study aimed exploring perceived applicability, usability of the system, learning and NLU performances during interviews with virtual patients. Participants were randomly assigned to Spoken or Typed condition. The results showed that VSP provides opportunity to practice medical interviewing skills before actual patient interaction. Results will be discussed with the demonstration of the system.

Literature

Barrows' (1993) definition of "standardized patients" (SP) both includes simulated patient and actual patient. While, simulated patient is a "person trained to simulate a patient's illness in a standardized way (p. 443), actual patient is "trained to present his or her own illness in a standardized way (p. 443).

Since then, SPs have been used for teaching and assessment in medical education (Adamo, 2004; Collins & Harden, 1999; Jack et al., 2009). Although SPs provide social skills, communication skills, judgment, and diagnostic acumen in real setting, they fall short of clinical interviewing skills (Talbot, Sagae, John, & Rizzo, 2012a). Moreover, despite the advantages of SPs, they are employed sparingly due to the high expense of hiring and training actors (Parsons et al., 2008). Also, although SPs will often provide verbally correct answers, the medical accuracy of the responses can be as low as 30% (Tamblyn et al., 2009).

Because of the limitations of SPs, there have been numerous attempts to replicate the experience on a computer for anytime-anywhere access to the experience and Virtual Patients have been started to be used instead of SPs. There have been variety of example applications called as Virtual Patients, including case presentations, interactive patient scenarios, virtual patient games, high fidelity software simulations, and virtual human conversational agents (Triola et al., 2006; Talbot, Sagae, John, & Rizzo, 2012b). Although, there have been different examples in the literature, intelligently designed simulations with reliable assessment needs to be designed and investigated. The purpose of this study was to design and explore the effectiveness of such Virtual standardized patients (VSP).

Methods

USC Standard Patient which provides engaging virtual standardized patient (VSP) enabling objective and meaningful assessment of learner after clinical virtual interview and diagnostic skills performance is used in the study. The system 1) Provides a conversational simulated patient used for medical training, 2) Capable of natural language interaction with verbal and nonverbal behavior responses 3) Offers consistent, objective experience and detailed user feedback to learners.

The purpose of the study was to examine the effects of two version of USC Standard Patient: Spoken versus Typed. Specifically, the study was designed to examine the following research questions:

- What is the perceived applicability of the participants?
- Is there a significant difference between typed and speech conditions in terms of usability of the system?
- Does two versions affect the type of questions asked, learning performances, and total turns asked during the medical interview with virtual standardized patients?

- Does two versions affect the NLU performance?

A total of 29 third-year medical students were recruited from Keck School of Medicine under the IRB exempt protocol and USAMRMC human protections approval. Participants were randomly assigned to either Spoken or Typed group. VSP were the same for both groups except the communication was either by voice or typing. All participants first watched the orientation video and completed the orientation survey. Afterwards, they interacted with two cases with two 5-minute attempts for each of them followed by viewing system feedback. The System Usability Scale (SUS) (Bangor, 2009) was administered for measuring usability of the system for several aspects, including effectiveness, efficiency, and satisfaction of use. Moreover, both descriptive and a repeated measure multivariate analysis of variance was conducted in SPSS Statistics v24. Also, qualitative analysis of the transcripts were done to assess natural language understanding (NLU) system performance.

Results

Participants' perceived applicability results showed that, majority of them in agreement that they can practice diagnostic interviewing skills. They also agree that it can prepare them to see both standardized patients and real patients. Participants were strongly agree, 22%, agree %29, somewhat agree 35%, and neither agree or disagree 14% to use the system given the change again.

SUS score for all of the participants were 75.8 which shows the system has very good usability. Moreover, analysis of the normalized SUS scores showed that two conditions, Typed and Speech are not statistically significant ($F(2, 28) = .016, p > .00$)

Results for Case 1: *Otisis Externa for Question asked, Performances, and Total turns* showed no significant main effects of Condition, and no significant multivariate effect of Attempt X Condition, but significant effect of Attempt. Results for Case 2: *Schizophrenia for Performance* showed significant main effects of Condition (Typed group showed higher performance than Speech group), significant effect of Attempt, but no significant multivariate effects of Attempt X Condition.

NLU Performance analysis showed that:

- Typed input is more accurate for NLU understanding
- Typed errors are uncommon and are not too likely to result in an NLU error
- Typed NLU improves upon second iteration of a case
- Verbal input is 15 points less accurate (consistent with prior studies)
- Verbal recognition errors are likely to result in NLU error

Discussion

Overall, participants liked the system and found it easy to operate. Participants in both conditions improved their learning performance from attempt one to two for both VSP cases. Performance improvements with the Typed and Speech groups showed similar pattern over time. Despite the overall learning gain, the results suggest that Typed group scored higher than Speech group. More accurate NLU understanding of Typed group and verbal recognition errors in NLU in Spoken group could have an affect the learning performances. However, participants perceived less sense of privacy w/ voice input.

References

- Adamo G. (2003). Simulated and standardized patients in OSCEs: achievements and challenges 1992–2003. *Med Teach*, 25, (262–70).
- Barrows H. (1993). An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med*, 68 (443–453).
- Collins, J. P. & Harden, R. M. (1999). The Use of Real Patients, Simulated Patients and Simulators in Clinical Examinations. *Med Teac*, 20, 508-521.
- Jack, B., Chetty, V., Anthony, D., Greenwald, J., Sanchez, G., Johnson, A., Forsythe, S., O'Donnell, J., Paasche-Orlow, M., Manasseh, C., Martin, S. & Culpepper, L. (2009) A Reengineered Hospital Discharge Program to Decrease Rehospitalization: A Randomized Trial. *Ann Intern Med*, 150(3), 178-187.
- Parsons, T. D., Kenny, P. et al. (2008). Objective Structured Clinical Interview Training using a Virtual Human Patient. *Studies of Health Informatics*, 132, 357-362.
- Talbot TB, Sagae K, John B, Rizzo AA. (2012a). Designing useful virtual standardized patient encounters. *Interservice/Industry Training, Simulation and Education Conference Proceedings*.
- Talbot TB, Sagae K, John B, Rizzo AA. (2012b). Sorting out the Virtual Patient: How to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer Mediated Simulation*, 4(3), 1-19.
- Tamblyn, R.M., Klass, D.J. & Schnabl, G.K. (2009). The accuracy of standardized patient presentation. *Medical Education*, 25(2), 100-109.
- Triola, M., Feldman, H., Kalet, A.L., Zabar, S., Kachur, E.K., Gillespie, C., Anderson, M., Griesser, C. & Lipkin, M. (2006). A Randomized Trial of Teaching Clinical Skills Using Virtual and Live Standardized Patients. *Journal of General Internal Medicine*, 21, 424-426.