Utilizing a Learning Strategy Analysis to Determine a System’s Potential Impact on Student Learning: The Augmented-Reality Emotive Virtual Patient System Platform

Sabrina Cervantes Villa	Scotty D. Craig	Djakhangir Zakhidov	Marjorie Zielke
Arizona State University	Arizona State University	The University of Texas at Dallas	The University of Texas at Dallas

A learning strategy analysis was performed on the Emotive Virtual Patient System, an augmented reality platform that teaches medical students doctor-patient communication skills. The Emotive Virtual Patient System is a complex mixed reality platform that includes both virtual and human peers/instructors who use natural language processing to provide feedback and dialog modeling as a means to improve patient communication learning outcomes. The learning strategy analysis (i.e., system learning strategy/component review, literature review, and heuristic evaluation of best practices) was conducted on the early system plans to determine its potential in supporting student learning and to provide short-and-long-term design considerations. The analysis identified three major categories for potential consideration: verbal interactions, user groups/system objective monitoring, and security. Specific recommendations were given for each of these areas, as supported by the literature.

INTRODUCTION

Augmented reality (AR) systems have become a popular topic of both exploration and study. Such systems allow users to interact with virtual objects in a physical space. One broad area of application for such conditions deals with training, where the user is attempting to learn a skill in a relatively low-stress environment. The current study analyzes an AR system with such a goal.

AR blends virtual realities with the real environment into one, single view (Zielke et al., 2018). AR systems can be complex with many interacting components. When applied for educational purposes, specific interactions between components may impact students’ learning. To understand the potential for learning and any opportunities for the current system, a heuristic evaluation method called learning strategy analysis (LSA) can be implemented (Roscoe et al., 2020). An LSA reviews a system's descriptions to identify the potential elements that are relevant to learning. It then provides a summary of the system’s potential as well as possible enhancements based on the existing research literature. Finally, it provides guidance in the form of design recommendations.

This method was applied to the Emotive Virtual Patient (EVP) System platform. The system includes an array of state-of-the-art technology developments. These include: a conversational virtual patient for practicing medical interviewing, automatic assessment of student performance, other virtual humans (VHs) developed for students to engage with social learning, and new interaction modes such as shared AR. It also includes a new user interface for students to select their preferred modality for social learning. The system’s key VH is an emotive virtual patient named Walter, designed for students to practice the communications component of the Objective Structured Clinical Examination (OSCE), clinical assessments for patient interactions.

Walter is a high-fidelity animated 3D character capable of natural conversations with users in addition to automatic assessment of student performance based on the standardized patient (SP) rubric. Walter expresses emotions through facial and body animations in response to questions. Walter can be accurately placed in a real-life examining room. In addition to Walter, the system includes a virtual professor, a virtual male peer, and a virtual female peer for students to learn from socially. These VHs can provide educational feedback to students and demonstrate how to do an OSCE interview. Students can select which of the VHs they want as collaborators for receiving feedback and for learning through observation. Finally, the system includes two novel modalities for collaborative interaction: distributed AR and co-located AR. Distributed AR enables an expert to remotely view a student’s performance during an OSCE and provide real-time feedback. Collaborative AR enables a colleague to observe a student’s performance from the same location and provide feedback. This AR research may be also be applicable for virtual reality and overall mixed reality systems.

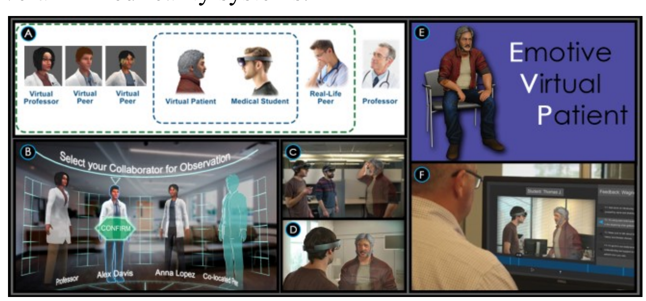

Figure 1. *Emotive Virtual Patient System*

Figure 1 offers an overview of how learners within the system collaborate. The interaction can involve a virtual professor, two virtual peers, the virtual patient, the medical student using the system, a real-life peer, and a real-life professor. When the application is launched, the user selects a collaborator for observation. The user can choose between the virtual professor character, the virtual male student, the virtual female student, or a real-life peer who can contribute to the conversation using co-located AR. In co-located AR, the user can conduct the interview of the virtual patient. A real-life peer can observe and give feedback on the interaction. If the user chooses a virtual character for observation, the user conducts the patient interview, and receives feedback afterwards from the selected virtual collaborator.

While the system presents a multi-disciplinary approach utilizing learning theory, social science and technology research additional potential design considerations may need to be considered. Using the initial system research proposal, the major features for analysis are summarized as: AR for educational purposes, pedagogical learning agents, natural language processing (NLP), and feedback with a parallel analysis of key social learning theories. Recommendations were then made regarding each of these features.

METHODS

LSA is a new heuristic evaluation method for learning technology that combines a review of the system to identify learning supports/scaffolds built into the system, literature review to identify best practices, and heuristic evaluation to determine overlap and potential improvements. This process has four steps: 1. identify system's learning strategies, 2. literature review on each strategy for effectiveness, 3. comparison between best practice and system, and 4. providing ratings and recommendations.

For the current LSA, relevant documents pertaining to the system was identified. This included the grant proposal and corresponding project summary. This led to identification of four learning supports and technologies used by the system as outlined above: 1. AR for educational purposes, 2. pedagogical learning agents, 3. NLP, and 4. feedback with a parallel analysis of key social learning theories. This was followed by literature reviews and verification of best practices. This process led to specific recommendations and potential rating impact levels. Table 1 includes the scale that was used to classify system ratings when identified. By explicitly considering each feature, the analysis could determine the system's potential for supporting student learning.

Table 1
Ratings of System Considerations

Rating	Description
1 (Low)	Minor items that may lead to unfavorable perception of the system by students
2 (Medium)	Considerable items that are likely to hinder students' understanding of the system
3 (High)	Major items that may disrupt students' ability to learn in or understand the system
4 (Critical)	Detrimental items that are extremely likely to hinder students' ability to learn

LITERATURE REVIEW RESULTS

Augmented Reality for Educational Purposes

There have been various degrees of application for AR in the field of education, with notable positive effects on student learning and motivation (e.g., Kamphuis et al., 2014). For medical instruction, AR has an extensive history of helping students visualize complex concepts (Kamphuis, 2014). It has also helped students train for intense medical situations (Conradi et al., 2009; Kamphuis, 2014).

This system utilizes AR alongside virtual humans. Effective virtual humans rely on establishing social presence (e.g., Mayer, 2014), and using nonverbal cue, dynamic dialog and modeling (Craig & Schroder, 2018). Social presence can be created via *embodiment cues*, for example. This includes the purposeful design of the agent's movement in the environment (Petersen, Mottleson & Makransky, 2021). The system aligns with the virtual human literature. Walter (Zielke et al., 2018) is a dynamic character that interacts in natural language provide interaction and models communication skills.

The literature emphasizes the use of structured interaction for computer-based collaboration (Dillenbourg & Fischer, 2007). This method of collaboration incorporates active interaction between students and virtual humans, which could utilize either a semi-structured script (focused dialogues within a set structure) or fully structured script for collaboration in the environment. These scripts are often chunked into micro collaborations that allow for focusing on unique elements and can be connected as needed for the learner. These scripts set forth a consistent role for students and pathways for the system to follow with the student. Novak and colleagues (2018) suggest that assigning students a role other than learner does not benefit overall learning. Thus, the system's current model of collaboration across experimental conditions is appropriate.

A final consideration for use of educational AR is accessibility. New learning technologies could have the most impact if they target the technology platforms that are accessible and familiar to the learners. This may mean that for a high degree of participants, their previous experience with new technologies (such as AR) can increase their general comfortability with the technology (Zielke et al., 2018).

While general experience with AR may have a relationship to socioeconomic status which could influence how their overall interaction with the system is interpreted (Ching et al., 2005), there is a growing trend to add AR into higher education learning environments which would make these interactions common place for learning (Becker et al., 2017). The EVP Platform, which utilizes state-of-the-art HoloLens II AR devices which currently cost about $4000 is integrated into the current learning environment and therefore accessibility will be controlled. Other forms of AR distribution could be considered in the future. For example, Siegle, Roscoe, Schroeder, and Craig (2020) suggest that immersive technologies for educational purposes may be developed for mid- to low ranges of devices that are increasingly more available and familiar. For example, the EVP system, or components of the system, could possibly be distributed over phone-based AR systems or other less expensive devices as they become available.

Pedagogical Agents for Learning

Overall, the system demonstrates effective use of virtual humans, in that the features of appearance, multiplicity, and effect of these agents align well with the literature (e.g., Craig & Schroeder, 2018; Kim & Baylor, 2016; Mayfield et al., 2019; Petersen et al., 2021; Schroeder & Craig, 2021).

The system supports interactions between medical students and three categories of virtual humans: a virtual professor, and two virtual peers (Zielke et al., 2018). The goals associated in interacting with the system's agents differ for the student, suggesting that the respective social learning theories behind the interactions differ as well. Overall, the system adheres to the best practice design principles for virtual humans related to using virtual humans as interactive dialogue systems, using them to support social presence in learning technologies, feedback support for learning, and providing modeling behaviors (Craig & Schroeder, 2018).

Bandura's (2001) social-cognitive theory argues three forms of agency: personal, proxy, and collective. In virtual environments, personal agency is directly affected by a student's interactions with their virtual peer. Kim and Baylor (2006) found that interactions with a virtual human whose level of perceived competency was being similar to the learner fostered higher levels of self-efficacy. The system aligns with this notion by implementing three distinct virtual humans with explicit roles.

Broadly, the system does seem to implement strong application of proxy agency via social modeling, which also increases self-efficacy in educational environments. Applications of collective agency in the system (via virtual humans) are also supported by the literature, in that student can interact with two virtual agents at once in certain experimental conditions (rather than being limited to one) (Baylor, 2009).

The interaction script should be developed for natural interaction. The literature emphasizes the importance of structured interaction between user and agent, in that the interaction is specifically designed to foster productive activities (e.g., Dillenbourg & Fischer, 2007; Vogel et al., 2017). It is recommended that the system utilizes a micro-collaboration script that are scripts which focus on dialogue and tasks that are both natural and that builds upon itself.

Feedback

Through use of both human and virtual peers/professors, the system's use of feedback is at the forefront of all interactions. Overall, literature pertaining to effective use of feedback is conflicting, leaving no single framework of what truly effective feedback looks like. Reviews of the literature support practices associated with *formative feedback* (Shute, 2008). This is feedback that benefits the learner.

Shute (2008) argues that formative feedback has two major modes: verification and elaboration. The system includes feedback from a remote instructor, sent directly into the EVP System App. Shute further suggests that this feedback *a) clearly identifies whether the student is right or wrong,* and *b) clearly and concisely offers the student response-specific elaboration*. Feedback that is perceived as long or complex can be frustrating or confusing. This type of feedback may be ignored or have a negative impact on student learning (Hattie & Timpreley, 2007; Shute, 2008). By offering concise feedback, it is more likely that the remote instructor will successfully help scaffold the information to the student (Hattie & Timperley, 2007).

Regarding feedback from peers, the literature suggests that structuring the way that feedback is offered, for example - how many points of criticism to point out (Patchan & Schunn, 2015), can be useful. In alignment with Piaget's concept of equally able partners, those peers who are of relatively similar levels of competence may especially benefit from such methods. The findings from the third iteration of this experiment may continue to build on this concept.

Natural Language Processing

The system utilizes NLP to emulate realistic dialogue between a student and patients/peers in the immersive environment. The literature suggests that ethical and effective use of NLP focus on four main areas: privacy; fairness, bias, and discrimination; abstraction and compartmentalization; and automation (Leidner & Plachouras, 2017).

Privacy is an important consideration for any system collecting user data for use of NLP and the literature suggests that a log-in system should be utilized to address this. Additionally, systems may benefit from explicit disclosure to the student regarding all personnel who will have access to their dialogue data. This information could possibly be included in a tutorial or navigation section prior to interactions.

The literature also suggests that a system with general NLP could consider the level of abstraction used, in that the system may not fully account for the full range of users (Leidner & Plachouras, 2017). For example, while some systems may account for English speakers of a certain dialect or accent, other students may have issues. This could lead to inaccurate interpretation of students' verbal input; which could lead to problems in how a student's performance is assessed and ultimately perceptions of system performance. This feeds into a larger issue of compartmentalization, focusing on small piece of the larger population, (Leidner & Plachouras, 2017). However, this is still an open research question because these items are based off broader concerns from the literature applied to educational NLP whereas the current EVP system is focused on teaching students patient communication skills, with an established assessment rubric which must be considered. Further, an interesting research question posed by this work is the degree that we can reconcile what is good patient communication and what a standardized patient (actor simulating a patient) or a real patient might reasonably understand.

Summary of Findings

Overall, the system performed very well in this evaluation and the system's design decisions for AR implementation appear to support learning. Specifically, the initial plans demonstrated apparent consideration and application of all forms of agency (Bandura, 2001) with the implementation of virtual humans. The system also had a good implementation for feedback that corresponded to recommendations (e.g., Hattie & Timpreley, 2007; Shute, 2008; Patchan & Schunn, 2015).

In summary, the learning strategies surrounding AR, virtual humans and feedback were found to be effectively implemented to support learning. With regards to NLP, there are always many considerations that could impact potential evaluation of specific users, and the EVP system has a unique purpose -- to teach patient communication techniques and to assess effectiveness. These dual objectives should be continually reviewed to determine if any enhancements to the NLP should be considered in future versions. These considerations were grouped into broader categories, verbal interactions, user group/system objective monitoring and security, which led to recommendations (see Table 2) that are discussed in the next section.

DISCUSSION/RECOMENDATIONS

LSA for a system is often performed early in the development process. The goal is to identify considerations with the system and to guide future steps in the research process i.e. developing the design document, finalizing iterative design decisions, focusing areas for monitoring, or guiding future system directions. The current LSA for the EVP system established three areas of consideration (AOC) including verbal interactions, user group/system objective monitoring, and security. These areas are explained below. See Table 2 for a summary of these recommendations and their rating levels.

The first AOC involves verbal interactions within the system. Generally, the literature supports structured or semi-structured scripts for collaborative interactions (Dillenbourg & Fischer, 2007). It also supports feedback that is formative and structured in nature, in that statements clearly verify how correct the learner is in their response and then elaborates on why or why not (Shute, 2008). The main goal of the EVP system is to develop conversation that combines NLP, visual environment, and nonverbal/verbal behavior of virtual human in the context of micro collaboration. It would be useful for the system to consider micro collaboration scripts, short feedback focused interactions between all agents. These scripts could be standardized to follow the patient communication rubric being taught. This would allow for interactive feedback and modeling of best practice. A rating level of 2 (medium) was assigned based on the potential impact to future iterations.

The second AOC deals with the continual reconciling of various user groups in light of the system's research objective – which is to teach patient communication techniques. While the system utilizes an advanced NLP system, the ability to interpret various English language accents or dialects should be monitored. Additionally, as the system matures, the literature suggests that displays of emotion range across cultures (Okur et al., 2018). This can affect the way that learners' performance is perceived by VHs. These ongoing performance findings will need to be considered in light of the system's objectives.

In addition, it is recommended that the AR technology introduction be continually be monitored to reflect the ability of the target population (Siegle et al., 2020). As stated, earlier AR technology is becoming more prevalent in higher education (Becker et al., 2017) and may not be completely novel to the current population (Zielke et al., 2018). This recommendation would be most useful as a monitoring goal. In that, it should be watched in the current system to determine the need in future versions. While these types of considerations might not be identified early on, monitoring should be implemented as different population test groups and situations emerge. The literature also suggests that user interface training would be helpful. This is especially important for users who may not have previous experience with AR technologies. As a result, a rating level of 2 (medium) was assigned because of the potential impact on the current system that could become more serious in the future and the need for ongoing monitoring.

Finally, system security, such as a log-in system, should be a critical consideration as the research is developed. These design constructs increase the system's level of privacy (Leidner & Plachouras, 2017). These recommendations are aspirational that can help guide future version of the system, as log-in options and system training would be a requirement for a full-scale system, while the current system is a research prototype. While all these recommendations may not be needed for the current EVP version, they are items that researchers should keep in mind in the future. Considering them early would prevent later redevelopment required for implementation. A rating level of 3 (high) was assigned to keep security considerations at the forefront as the system is further developed.

Table 2
Summary of Suggested Recommendations

Rating	AOC	Suggested Recommendation
2 (Medium)	Verbal interactions	Consider micro-collaboration scripts between all entities. Consider standardized of scripts and feedback.
2 (Medium)	User group/ system objective monitoring	Monitor accents cultural differences in expressing emotion, and how this will influence assessment of OSCE in light of patient communication learning objectives. Consider how inexperienced users of AR will interpret and interact with the system.
3 (High)	Security	Offer students a log-in option to the system. Implement a tutorial of interface navigation.

System Development Update

The current analysis was performed on the initial plan for the learning strategies of the system. Partially based on these recommendations and on best practices, many of these suggestions have been addressed. For example, the current system does have a log in system and data security protocol. Accents have been monitored and do not seem to be an issue with the current NLP/system objectives at this time. The NLP is constantly under review and we currently posit that our overall system approach to the EVP platform may enhance performance. The evaluation rubric and evaluation of what would be realistic standards for real humans versus virtual humans are important considerations to keep in mind.

Conclusion

To determine the potential for supporting student learning with the EVP system (Zielke et al., 2018), an LSA was conducted. The analysis demonstrates how a heuristic evaluation method can be applied to review a system's impact on future users early in the development process. The LSA indicated that the EVP system was well aligned with best practices with three potential future consideration items. The LSA items provide areas for monitoring going forward and could serve to inform the system research agenda.

The current paper was supported by funding from the National Science Foundation Grant No.1917994. The statements within this work do not reflect the opinion, view, or policy of the funding agency.

References

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26.

Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364* (1535), 3559-3565. https://doi.org/10.1098/rstb.2009.0148

Becker, S. A., Cummins, M., Davis, A., Freeman, A., Hall, C. G., & Ananthanarayanan, V. (2017). NMC horizon report: 2017 higher education edition (pp. 1-60). The New Media Consortium.

Ching, C. C., Basham, J. D., & Jang, E. (2005). The legacy of the digital divide: Gender, socioeconomic status, and early exposure as predictors of full-spectrum technology use among young adults. *Urban Education, 40(4)*, 394-411.

Conradi, E., Kavia, S., Burden, D., Rice, A., Woodham, L., Beaumont, C., Savin-Baden, M., & Poulton, T. (2009). Virtual patients in a virtual world: Training paramedic students for practice. Medical Teacher, 31(8), 713-720. https://doi.org/10.1080/01421590903134160

Craig, S. D., & Schroeder, N. L. (2018). Design principles for virtual humans in educational technology environments. In *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (pp. 128-139). Taylor and Francis. https://doi.org/10.4324/9781315109503

Dillenbourg, P. & Fischer, F. (2007). Basics of computer-supported collaborative learning. *Zeitschrift für Berufsund Wirtschaftspädagogik*. 21, 111-130. https://doi.org/10.1007/978-1-4020-9827-7_1

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112. https://doi.org/10.3389/fpsyg.2019.03087

Kamphuis, C., Barsom, E., Schijven, M., & Christoph, N. (2014). Augmented reality in medical education? *Perspectives on Medical Education*, *3*(4), 300-311. https://doi.org/10.1007/s40037-013-0107-7

Kim, Y., & Baylor, A. L. (2006). A social-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development, 54*(6), 569-596. https://psycnet.apa.org/doi/10.1007/s11423-006-0637-3

Kim, Y., & Baylor, A. L. (2016). Research-based design of pedagogical agent roles: A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education, 26*(1), 160-169. http://dx.doi.org/10.1007%2Fs40593-015-0055-y

Leidner, J. L., & Plachouras, V. (2017). Ethical by design: Ethics best practices for natural language processing. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, 30-40. https://www.aclweb.org/anthology/W17-1604.pdf

Mayer, R. E. (2014). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *Cambridge handbooks in psychology. The Cambridge handbook of multimedia learning* (p. 345–368). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.017

Mayfield, E., Madaio, M., Prabhumoye, S., Gerritsen, D., McLaughlin, B., Dixon-Román, E., & Black, A. W. (2019). Equity beyond bias in language technologies for education. In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, 444-460. http://dx.doi.org/10.18653/v1/W19-4446

Novak, K., King, A., & Roscoe, R. D. (2018) Exploring higher agency roles for learning with educational technology and multimedia. https://doi.org/10.1177%2F1541931218621468

Okur, E., Aslan, S., Alyuz, N., Esme, A. A., & Baker, R. S. (2018). Role of socio-cultural differences in labeling students' affective states. In *International Conference on Artificial Intelligence in Education*. 367-380. Springer, Cham. https://doi.org/10.1007/978-3-319-93843-1_27

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science*, *43*(5), 591-614. https://doi.org/10.1007/s11251-015-9353-x

Petersen, G. B., Mottelson, A., & Makransky, G. (2021). Pedagogical Agents in educational VR: An in the wild study. In *CHI Conference on Human Factors in Computing Systems (CHI '21)*, Yoko-hama, Japan. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3445760

Roscoe, R. D., McNicol, S., Raghav Bhat, K., & Craig, S. D. (2020). A heuristic evaluative framework for self-regulated learning design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, (64)*1, 1775-1779. Sage CA: Los Angeles, CA. https://doi.org/10.1177%2F1071181320641429

Schroeder, N. L., & Craig, S. D. (2021). Learning with virtual humans: Introduction to the special issue, *Journal of Research on Technology in Education,* 53, 1-7. https://doi.org/10.1080/15391523.2020.1863114

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research, 78*(1), 153-189. https://doi.org/10.3102%2F0034654307313795

Siegle, R. F., Roscoe, R. D., Schroeder, N. L., & Craig, S. D. (2020, December). Immersive Learning Environments at Scale: Constraints and Opportunities. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64, No. 1, pp. 1165-1169). Sage CA: Los Angeles, CA: SAGE Publications.

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review*, *29*(3), 477-511. https://eric.ed.gov/?id=EJ1150773

Zielke, M. A., Zakhidov, D., Hardee, G. M., Pradeep, J., Evans, L., Lodhi, Z., & Ward, E. (2018). Exploring medical cyberlearning for work at the human/technology frontier with the mixed-reality emotive virtual human system platform. In *2018 IEEE 6th international conference on serious games and applications for health (segah)* (pp. 1-8). IEEE. https://doi-org.ezproxy1.lib.asu.edu/10.1109/SeGAH.2018.8401366