Assessing Pediatric Life Support Skills Using Augmented Reality Medical Simulation With Eye Tracking: A Pilot Study

Jimmy Qian, BS
Asheen Rama, MD
Ellen Wang, MD

Tammy Wang, MD
Olivia Hess, BS
Michael Khoury, BS

Christian Jackson, MS
Thomas J. Caruso, MD

Introduction

Augmented reality (AR) is an emerging adjunct for medical simulation. Distinct from virtual reality during which users are completely immersed in a computer-generated environment via headset, AR uses head-mounted devices that overlay holographic projections onto the real world. As a result, users can simultaneously interact with people and holographic objects, reducing reliance on physical medical simulation equipment. This technology has gained traction in recent years to advance immersive clinical learning experiences.[1-3]

Another developing technology in medical education is eye tracking, which relies on devices that record moment-to-moment gaze patterns. Eye tracking has been used during medical simulation, resulting in promising developments in clinical training and assessment such as data-driven feedback for trainees on their technical skills or situation awareness.[4-10] Eye tracking can help instruct learners' situational awareness based on where their attention is visually directed. Eye tracking involves the use of near infrared light, reflections from the users' eyes, and cameras to determine where the users are fixating their gazes. Although dedicated eye tracking headwear provide industry standard tracking, in some augmented reality headsets, such as the Magic Leap 1 (ML1, Magic Leap Incorporation, Plantation, FL), user gaze patterns are tracked to increase the fidelity of the holographic experience. The ML1

eye tracking combines short exposure (200 frames per second) with long exposure (30-60 frames per second) images to track eye movement.[11] Eye gaze information is provided for every rendered frame.[11] The setup is binocular and is always worldview, thus it tracks and predicts eye direction even if the eye direction falls outside the ML1's render field of view.[11]

To combine the use of AR simulation and the assessment capabilities of eye tracking, an AR simulator with gaze tracking has been developed, hosted on the ML1 headset.[11] The simulator software (Chariot Augmented Reality Medical [CHARM] simulator, Stanford Chariot Program, Stanford, CA) is a multiplayer, in-person or distance experience within a mixed reality world of holographic and natural assets. Holographic patients, beds, and monitors are projected onto the real world and modulated in real time by holographic controllers seen only by simulation instructors. Simulations are not coded with predetermined sequences, but rather adjusted in response to participant performance. This novel approach of combining eye tracking in AR simulation could enhance the development of behavioral skills by providing integrated, real-time learning opportunities triggered by gaze behavior. The ability for remote participation provides an opportunity for instructors to teach participants in any area of the world given internet and ML1 headsets.

We examined participant interaction during a pediatric cardiac arrest simulation using the novel AR software with integrated gaze tracking. The goal was to determine whether AR simulation could be used to meet goals commonly targeted in intraoperative crisis management simulations, such as recognition of shockable rhythms with appropriate treatments. The primary aim assessed performance using a validated scoring instrument (the Anesthesia-centric Pediatric Advanced Life Support [A-PALS] tool) during a standardized AR simulation. Our primary hypothesis was that the composite faculty score would be higher than the composite score of residents and fellows. Secondary aims explored recognition of shockable rhythm via gaze patterns and participant satisfaction.

Methods

Context

This project was conducted at an academic, pediatric hospital in Northern California between March and May 2021. The Stanford University Institutional Review Board provided an exemption because the medical simulations were performed as part of an ongoing educational program that is provided to anesthesiology trainees during their pediatric anesthesiology rotation and those attendings who elect to participate. Post-graduate year (PGY) 3 and 4 anesthesiology residents on their pediatric anesthesiology rotation, pediatric anesthesiology fellows, and pediatric

anesthesiology attending physicians participated in a pediatric cardiac arrest simulation using AR. Participants with glasses were excluded given that the AR headsets do not comfortably fit over glasses since prescription inserts for the headsets were not available. Individuals with nausea, severe motion sickness, or seizures were also excluded.

Intervention

Prior to each AR simulation, participants completed a brief questionnaire that provided demographic information, stage of clinical training, AR experience, and advanced life support experience (for example, if participants had American Heart Association certifications in Pediatric Advanced Life Support) (Table 1). Participants wore a ML1 headset installed with the CHARM simulator software.[11] The simulation equipment included a real pediatric chest task-trainer on a table. Through the AR headset, holograms were overlaid onto the real chest trainer, providing participants with a view of a full pediatric holographic patient in a holographic hospital bed with a nearby, holographic monitor displaying heart rate, blood pressure, plethysmograph, oxygen saturation, electrocardiogram (ECG), and end-tidal carbon dioxide (Figure 1). Simulations were video recorded.

After the participants completed the baseline questionnaire on REDCap,[12,13] an attending anesthesiologist, who served as the simulation instructor for all simulations, oriented the participants. Participants learned about the similarities and differences of traditional simulation and AR simulation, discussed examples of prior simulation experiences, and became familiarized with the equipment. The instructor outlined the learning goals for the simulation, including appropriate management of a pediatric cardiac arrest using closed-loop communication with initial responders. Resident and fellow performance was not included in their formal rotation evaluation.

Each participant assumed the role of the code leader for the simulation. For each simulation, the same 2 confederate research associates assisted the resuscitation efforts; one played the compressor role and the other played the role of a PGY 2 anesthesia resident. As the code leader, the participant verbalized changes in the patient conditions, such as relevant vital signs. Participants and the 2 confederates used closed-loop communication to verbalize requests and subsequent actions. For example, the compressor verbalized the rate of compressions they planned to administer while compressing the chest task trainer. Throughout the simulation, the AR headset tracked participant eye gaze. Participants were not aware of the gaze tracking to optimize natural gaze patterns and minimize the risk of participants altering their gaze patterns.

The simulation scenario began with a 5-year-old hyperkalemic male with Trisomy 21 and B-cell acute lymphoblastic leukemia who was undergoing general anesthesia for a peripherally inserted central catheter placement, lumbar puncture, and bone marrow aspiration. Shortly after the simulation began, the holographic patient developed signs of hyperkalemia progressing to ventricular tachycardia and fibrillation. The simulation progression and assessment of participant competency used the A-PALS scoring instrument.[14] The clinical scenario was standardized across participants such that verbalized actions elicited the same response from the confederates and holograms per the simulation protocol.[14] Based on verbal cues from participants, the simulation instructor adjusted the patient's holographic vital signs in real time while the confederates adjusted their responses to the participant's questions and requests for information. After an average of 8 minutes, the simulation ended when ventricular fibrillation was managed with return of spontaneous circulation and disposition was discussed.

After the simulation, participants completed the Simulation Design Scale.[15,16] The simulation instructor debriefed with the participants to discuss appropriate management of the simulated scenario and opportunities for improvement.

Outcomes

The primary outcome explored clinician performance, stratified by level of training.[14] One secondary outcome explored the use of eye tracking as a measure of shockable rhythm recognition, stratified by level of training and performance. Another secondary outcome explored participant satisfaction with the AR simulation.[15,16]

Measures

The primary outcome of clinician performance was assessed using the A-PALS scoring instrument (Appendix A).[14] Video recordings of simulations were reviewed to assign A-PALS scores for each participant, following the scoring guidelines. For consistency, the same study investigator scored all simulations and was blinded to the level of the participant. The A-PALS score was reported in aggregate and by section (general assessment, crisis-specific assessment, hyperkalemia-specific assessment, ventricular fibrillation management, and return of spontaneous circulation management).

Scores were reported by level of training, including resident, fellow, and attending. Scores were reported as percentage of tasks appropriately performed per A-PALS scoring (Appendix A). We hypothesized that attendings would perform better than residents and fellows. Given a reference standard of 82.6% ± 5.0% score for strong performances during a pediatric hyperkalemic arrest and a reference standard of 67.2% ± 3.8% A-PALS score for poor performances, a power of 80%, and alpha level of 0.05, 3 participants per group were needed to demonstrate performance differences among these 3 groups, for a total of 9 participants.[14] With 9 pediatric anesthesiology fellows in the fellowship program, we targeted recruitment of a balanced number of 9 residents and 9 faculty to match the target of 9 fellows. Residents were recruited sequentially as they were assigned to rotate on pediatric anesthesiology. Faculty were solicited through verbal and electronic communications. All participation was voluntary.

The secondary outcomes were exploratory. The first secondary outcome was defined as the amount of time that elapsed between the onset of ventricular fibrillation or ventricular tachycardia and when the participant gazed at the holographic ECG. It was calculated by using eye tracking

data that provided a timestamp of when the participant directly looked at the ECG, which was also confirmed by verbal cues the participant gave after recognizing the irregular rhythm. The gaze timestamp was compared to when the shockable rhythm first began, which was also recorded by the software. We required a verbal recognition of the rhythm to accompany the gaze given previous distinctions between random gazes and gazes that denote recognition.[17] The gaze patterns, time stamps, and rhythm logs were automatically tracked by the CHARM software and downloaded as comma-separated values (.csv) files at the conclusion of the simulation. Gaze patterns were stratified by resident, fellow, and attending, and also by performance based on A-PALS scores.

The final secondary outcome assessed simulation satisfaction using SDS scores that were calculated by applying the scoring algorithms to the participant questionnaire.[15,16] The SDS survey was originally developed to assess participants' satisfaction with multiple facets of a simulation. It was used in this study to determine if the AR stimulation satisfaction scores rated well in typical domains of in-person simulation, including categories such as fidelity and problem solving (Appendix B).

Analysis

We calculated the mean A-PALS score and mean time to recognize the shockable rhythm for all participants. A-PALS score, time to shockable rhythm recognition, and SDS score were stratified by training level, and a one-way analysis of variance (ANOVA) was used to determine whether there were any group differences for the A-PALS scores. Participants were placed into high, medium, and low performing groups based on the tercile cutoffs of the A-PALS scores. Results were considered significant using a *P* value of .05.

RESULTS

Participants

We enrolled 27 participants with 9 each of residents, fellows, and attending physicians. Many participants were experienced in providing resuscitation and were trained in communication or life support skills, while over half had prior exposure to AR (Table 1).

Performance

Regarding A-PALS scores, residents, fellows, and attendings averaged 69.7% (±10.4), 66.1% (±10.7), and 70.2% (±11.9), respectively (*P* = .692) (Figure 2). For general assessment, residents scored similarly to fellows and attendings with averages of 88.9% (±22.0), 96.3% (±11.1), and 94.4% (±16.7), respectively (*P* = .642). For hyperkalemia-specific assessment, residents, fellows, and attendings also scored similarly, with average scores of 49.2% (±21.3), 49.2% (±24.7), and 52.8% (±19.5), respectively (*P* = .925). For ventricular fibrillation management, residents and fellows averaged 35.6% (±4.2) and 35.7% (±7.6) respectively compared to 30.3% (±7.0) for attendings (*P* = .149). For return of spontaneous circulation management, residents and fellows scored similar to attendings with averages of 29.6% (±11.1), 33.3% (±0.0), and 48.1% (±29.4), respectively (*P* = 0.093). A-PALS overall and section scores did not significantly differ between residents, fellows, and attendings (Figure 2).

Eye Tracking Gaze and Rhythm Recognition

For 15 of the 27 participants, the eye tracking data were confirmed by the participants' verbal signal that they observed the rhythm and their resulting management. Of the 15 participants whose eye tracking data were confirmed, there were 5 residents, 5 fellows, and 5 attendings. On average, residents, fellows, and attendings took 9.6 (±7.6), 1.8 (±1.9), and 7.3 (±9.2) seconds, respectively, to gaze at ventricular tachycardia after it commenced and 5.5 (±9.3), 9.4 (±11.3), and 3.1 (±1.3) seconds, respectively, to gaze at ventricular fibrillation after it commenced.

We also clustered participants based on their A-PALS overall scores into groups of low, medium and high scorers, which had 6, 4, and 5 participants, respectively. We considered an overall A-PALS score to be low between 48.3 and 62.8, medium between 62.8 and 76.0, and high between 76.0 and 85.7, based on terciles. On average, low, medium, and high scorers took 11.6 (±10.6), 1.9 (±1.5), and 6.1 (±2.6) seconds, respectively, to gaze at ventricular tachycardia and 3.1 (±2.3), 3.5 (±3.4), and 11.5 (±13.0) seconds, respectively, to gaze at ventricular fibrillation.

Satisfaction

Assessing satisfaction using the SDS, all participants rated the simulation highly (Figure 3). Almost all satisfaction scores in each category from each cohort of participants fell between 4 and 5. There were no differences in satisfaction responses between groups.

DISCUSSION

We demonstrated the feasibility of using AR with eye tracking to conduct simulations assessing a hyperkalemic arrest treatment algorithm across residents, fellows, and attendings. With this innovative approach, we accomplished intraoperative crisis management simulation goals of pattern recognition and treatment without the typical equipment needed for high-fidelity simulation. A-PALS scores demonstrated successful progression through the simulation, with no differences across training levels. The average A-PALS scores were similar to benchmark scores achieved with in-person simulations.[14] Integrated eye tracking provided insight into shockable rhythm recognition across training levels, although these results require more investigation with greater power to discern implications. Simulation satisfaction was high using this novel technology.

The primary outcome demonstrated that the simulation was achievable, given how easy and low-cost the setup was. The simulation required 2 ML1 headsets and software, a chest task trainer, and a table, without the need for additional equipment or monitors. Although a task trainer was used in this simulation, the software has been used without any trainers, relying on the holographic images alone. This study illustrates how AR can expand pediatric crisis simulation and education into settings where there is limited access to simulation equipment. AR headsets such as the ML1 used in our simulation cost approximately $1,800 apiece (we used 2 at a cost of $3,600), which is substantially lower than traditional simulation manikins and affiliated equipment which can cost over $100,000 USD.[18] In addition, the software has multi-player, distance-

enabled capabilities, allowing instructors and trainees to simultaneously participate around a holographic simulation while not physically located in the same space. This type of simulation may enable simulations to be performed in limited resourced settings with remote instructors, while providing advanced simulation feedback such as the time until recognition of a shockable rhythm.

We successfully used integrated eye tracking during the AR simulation without a separate, dedicated eye tracking device, and there were no meaningful differences across residents, fellows, and attendings. Because participants were not aware of the gaze tracking capabilities, it is unlikely that their patterns were implicitly or explicitly altered. However, given the lack of statistical power, these results should be interpreted with caution. Even though the gaze patterns were not predictive of performance in our trial, level of training has previously been correlated with successful pediatric resuscitation.[19,20] Given that the A-PALS score did not differ across the 3 training levels, the difficulty of the simulation may have lacked discriminative properties to differentiate skills.

Eye tracking has been used in medical education for clinical training, assessment, and feedback.[7,8] However, eye tracking integrated into an AR simulator is a novel approach.[11] Previous investigators have explored gaze patterns during traditional simulations. For example, junior and expert surgeons had differences in gaze behavior during open inguinal hernia repair,[21] while pediatric intensive care physicians' gaze patterns differed compared to pediatric trainees, general pediatricians, and pediatric emergency medicine physicians during ventricular fibrillation simulations.[6] Situational awareness between level of training was explored through frequency and gaze duration at mannikin chest and airway compared to attention to the posted algorithm and defibrillator.[6] Even though there were no differences in gaze patterns during our study, it does highlight the question of what gaze patterns may be most useful in the context of pediatric crisis management assessment and training. Embedding eye-tracking within

AR provides copious data that may be used to discover the most meaningful gaze patterns, including gaze location over time throughout the simulation, dwell times, fixation counts, and number of gazes at the monitor.[22] It is possible that more nuanced criteria regarding gaze patterns may be necessary to detect differences between practitioners.

We also found no difference in satisfaction, stratified by level and performance, suggesting that the novel simulation medium was an acceptable tool. Poor performance did not appear to be attributed to the novelty of the AR technology. Consistently high satisfaction across training levels is promising because it suggests that this technologically advanced tool is equally satisfying to personnel of different ages and backgrounds. Residents, fellows, and attendings had the highest mean rating for the *feedback* category of the SDS, which encompassed the opportunity to engage with instructors and the use of analyzing one's own actions and behavior during the simulation.

There were several limitations. Given the single site recruitment, it will be necessary to collect data in other settings to ensure generalizability. It will also be helpful to conduct controlled trials with larger sample sizes. Also, we attempted to score the participants with a rater blinded to position of training. It is possible that the rater may have hypothesized a participant's level of training based on perceived age. However, the rater was unaware of the goals of the study and given the insignificant differences between groups, it is unlikely that the rater provided biased assessments. The AR headsets were not compatible with corrective lenses, which are used by many physicians, so further study is needed in those populations. Additionally, we were limited by the types of eye tracking we could use within the CHARM simulator, prohibiting our ability to investigate patterns across gaze, fixation, and saccades at the level of detail that would be available with a dedicated eye tracking device. The inability to register a participant's peripheral vision is a limitation specific to the eye tracking software and may limit our assessment of gaze patterns in some individuals. Because we required verbal confirmation to accompany the gaze,

not every participant's gaze data were included. The concept that a gaze denotes understanding is a debated topic in the field of eye tracking in medical education. While most agree that gaze patterns may be helpful, it is important not to overstate the implications of gaze patterns, as a gaze may not necessarily imply recognition and comprehension. We chose a conservative approach to the analysis of our gaze patterns at the expense of statistical power. Finally, medical simulations might not reflect real clinical practice, and therefore any study using simulations as endpoints may be fundamentally constrained.

In conclusion, we presented a novel approach applying eye tracking within an AR simulator to assess 3 levels of anesthesiologists during an AR-enabled pediatric cardiac arrest. Using a validated scoring system, we were able to successfully progress participants through the simulation as demonstrated by typical A-PALS scores. We also collected gaze patterns and received high satisfaction from all the participants. While the agreement between A-PALS score and gaze patterns is promising, further research is needed to fully demonstrate the use of AR eye tracking for medical training and assessment. Though eye tracking within AR medical simulation is in the early stages of adoption, physicians of multiple training levels were satisfied with the technology.

References

1. Tang KS, Cheng DL, Mi E, Greenberg PB. Augmented reality in medical education: a systematic review. *Can Med Educ J.* 2020 Mar 16;11(1):e81-e96.

2. Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. *Surg Endosc.* 2016 Oct;30(10):4174-83.

3. Toto RL, Vorel ES, Khoon-Yen ET, et al. Augmented reality in pediatric septic shock simulation: randomized controlled feasibility trial. *JMIR Med Educ.* 2021;7(4):e29899.

4. Chen H-E, Bhide RR, Pepley DF, et al. Can eye tracking be used to predict performance improvements in simulated medical training? A case study in central venous catheterization. *Proc Int Symp Hum Factors Ergon Healthc.* 2019;8(1):110-4.

5. Tahri Sqalli M, Al-Thani D, Elshazly M, Al-Hijji

6. Katz TA, Weinberg DD, Fishman CE, et al. Visual attention on a respiratory function monitor during simulated neonatal resuscitation: an eye-

continued from previous page

tracking study. *Arch Dis Child. Fetal Neonatal Ed.* 2019;104:F259-F264.

7. Tien T, Pucher PH, Sodergren MH, et al. Eye tracking for skills assessment and training: a systematic review. *J Surg Res.* 2014;191(1):169-78.

8. Ashraf H, Sodergren MH, Merali N, et al. Eye-tracking technology in medical education: a systematic review. *Med Teach.* 2018;40(1):62-9.

9. Desvergez A, Winer A, Gouyon JB, Descoins M. An observational study using eye tracking to assess resident and senior anesthetists' situation awareness and visual perception in postpartum hemorrhage high fidelity simulation. *PLoS One.* 2019;14(8):e0221515.

10. Krois W, Reck-Burneo CA, Gröpel P, et al. Joint attention in a laparoscopic simulation-based training: a pilot study on camera work, gaze behavior, and surgical performance in laparoscopic surgery. *J Laparoendosc Adv Surg Tech A.* 2020;30(5):564-8.

11. Caruso T, Hess O, Roy K, et al. Integrated eye tracking on Magic Leap One during augmented reality medical simulation: a technical report. *BMJ Simul and Technol Enhanc Learn.* 2021;7(5):431-4.

12. PA Harris, R Taylor, R Thielke, et al. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-81.

13. PA Harris, R Taylor, BL Minor, et al. The REDCap consortium: building an international community of software partners. *J Biomed Inform.* 2019;95:103208.

14. Watkins SC, Nietert PJ, Hughes E, et al. Assessment tools for use during anesthesia-centric pediatric advanced life support training and evaluation. *Am J Med Sci.* 2017;353(6):516-22.

15. Smith SJ, Roehrs CJ. High-fidelity simulation: factors correlated with nursing student satisfaction and self-confidence. *Nurs Educ Perspect.* 2009;30(2):74-8.

16. Fountain RA, Alfred D. Student satisfaction with high-fidelity simulation: does it correlate with learning styles? *Nurs Educ Perspect.* 2009;30(2):96-8.

17. Stiegler MP, Gaba DM. Eye tracking to acquire insight into the cognitive processes of clinicians: is "looking" the same as "seeing"? *Simul Healthc.* 2015;10(5):329-30.

18. Lapkin S, Levett-Jones T. A cost-utility analysis of medium vs. high-fidelity human patient simulation manikins in nursing education. *J Clin Nurs.* 2011;20(23-24):3543-52.

19. Brett-Fleegler MB, Vinci RJ, Weiner DL, et al. A simulator-based tool that assesses pediatric resident resuscitation competency. *Pediatrics.* 2008;121(3):e597-e603.

20. Donoghue A, Nishisaki A, Sutton R, Hales R, Boulet J. Reliability and validity of a scoring instrument for clinical performance during Pediatric Advanced Life Support simulation scenarios. *Resuscitation.* 2010;81(3):331-6.

21. Tien T, Pucher PH, Sodergren MH, et al. Differences in gaze behaviour of expert and junior surgeons performing open inguinal hernia repair. *Surg Endosc. 2015*;29(2):405-13.

22. Roland D. What are you looking at? *Arch Dis Child.* 2018;103(12):1098-9.

The following authors are at Stanford School of Medicine, Stanford, CA: **Jimmy Qian** *is a Medical Student;* **Olivia Hess** *is a medical student. The following authors are in the Department of Anesthesiology, Perioperative and Pain Medicine, Division of Pediatric Anesthesiology, at Stanford University School of Medicine, Stanford, CA:* **Asheen Rama** *is a Clinical Assistant Professor;* **Ellen Wang** *is a Clinical Associate Professor;* **Tammy Wang** *is a Clinical Associate Professor;* **Christian Jackson** *is a Statistician;* **Thomas J. Caruso** *is a Clinical Professor.* **Michael Khoury** *is a Research Assistant in the Stanford Chariot Program at Stanford School of Medicine, Stanford, CA and at Lucile Packard Children's Hospital Stanford, Palo Alto, CA.*

Corresponding author: *Thomas J. Caruso, MD, Department of Anesthesiology, Perioperative and Pain Medicine, Division of Pediatric Anesthesiology, Stanford University School of Medicine, 453 Quarry Road, MC 5663, Palo Alto, CA, 94304*

Email address: *Thomas J. Caruso: tjcaruso@stanford.edu*

Abstract

Background: Augmented reality (AR) and eye tracking are promising adjuncts for medical simulation, but they have remained distinct tools. The recently developed Chariot Augmented Reality Medical (CHARM) Simulator combines AR medical simulation with eye tracking. We present a novel approach to applying eye tracking within an AR simulation to assess anesthesiologists during an AR pediatric life support simulation. The primary aim was to explore clinician performance in the simulation. Secondary outcomes explored eye tracking as a measure of shockable rhythm recognition and participant satisfaction.

Methods: Anesthesiology residents, pediatric anesthesiology fellows, and attending pediatric anesthesiologists were recruited. Using CHARM, they participated in a pediatric crisis simulation. Performance was scored using the Anesthesia-centric Pediatric Advanced Life Support (A-PALS) scoring instrument, and eye tracking data were analyzed. The Simulation Design Scale measured participant satisfaction.

Results: Nine each of residents, fellows, and attendings participated for a total of 27. We were able to successfully progress participants through the AR simulation as demonstrated by typical A-PALS performance scores. We observed no differences in performance across training levels. Eye tracking data successfully allowed comparisons of time to rhythm recognition across training levels, revealing no differences. Finally, simulation satisfaction was high across all participants.

Conclusions: While the agreement between A-PALS score and gaze patterns is promising, further research is needed to fully demonstrate the use of AR eye tracking for medical training and assessment. Physicians of multiple training levels were satisfied with the technology.

Keywords: Augmented reality, eye tracking, medical simulation, pediatric advanced life support

Figures

Figure 1. CHARM simulator holographic view, shown from the vantage point of the simulation controller and narrator.
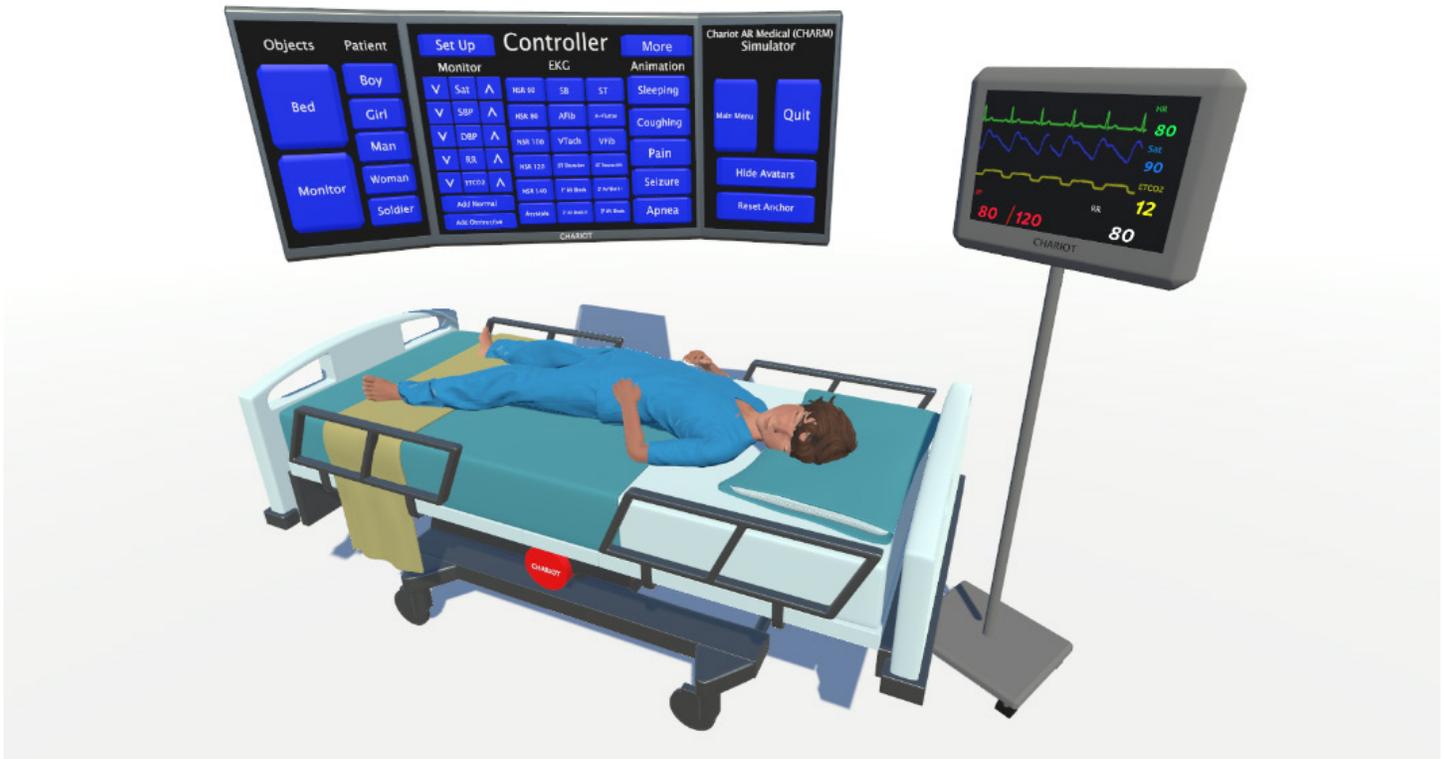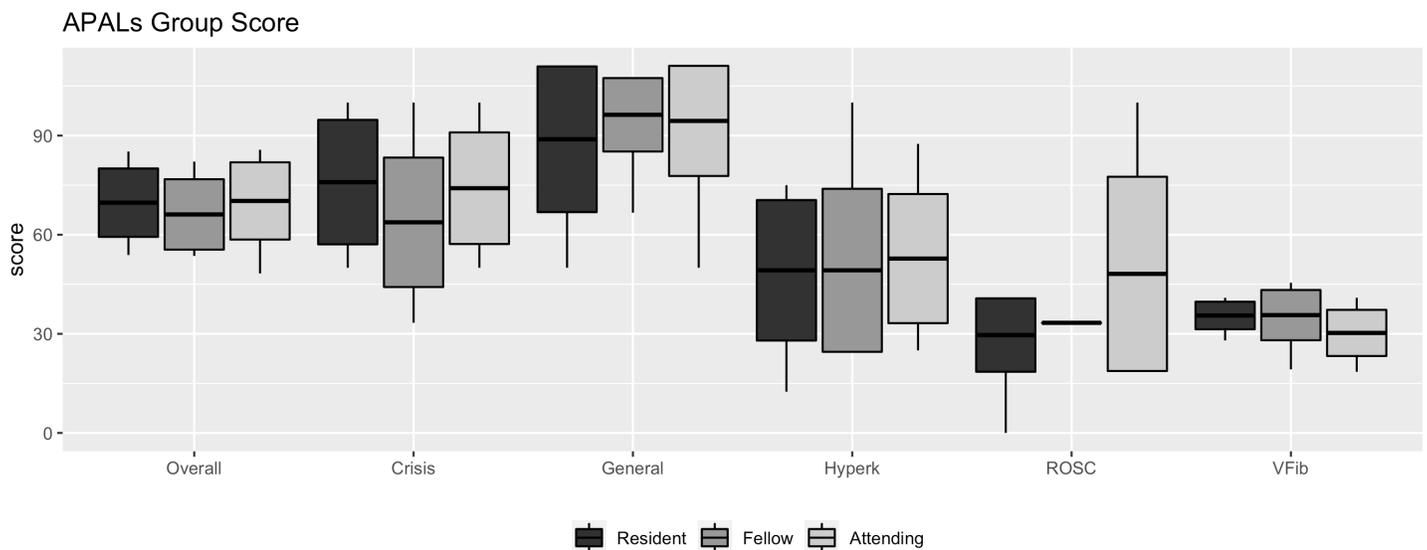

Figure 2. Mean and standard deviations of A-PALS scores. The overall A-PALS score, as well as the individual subscores, are included. Mean scores are grouped by training level.
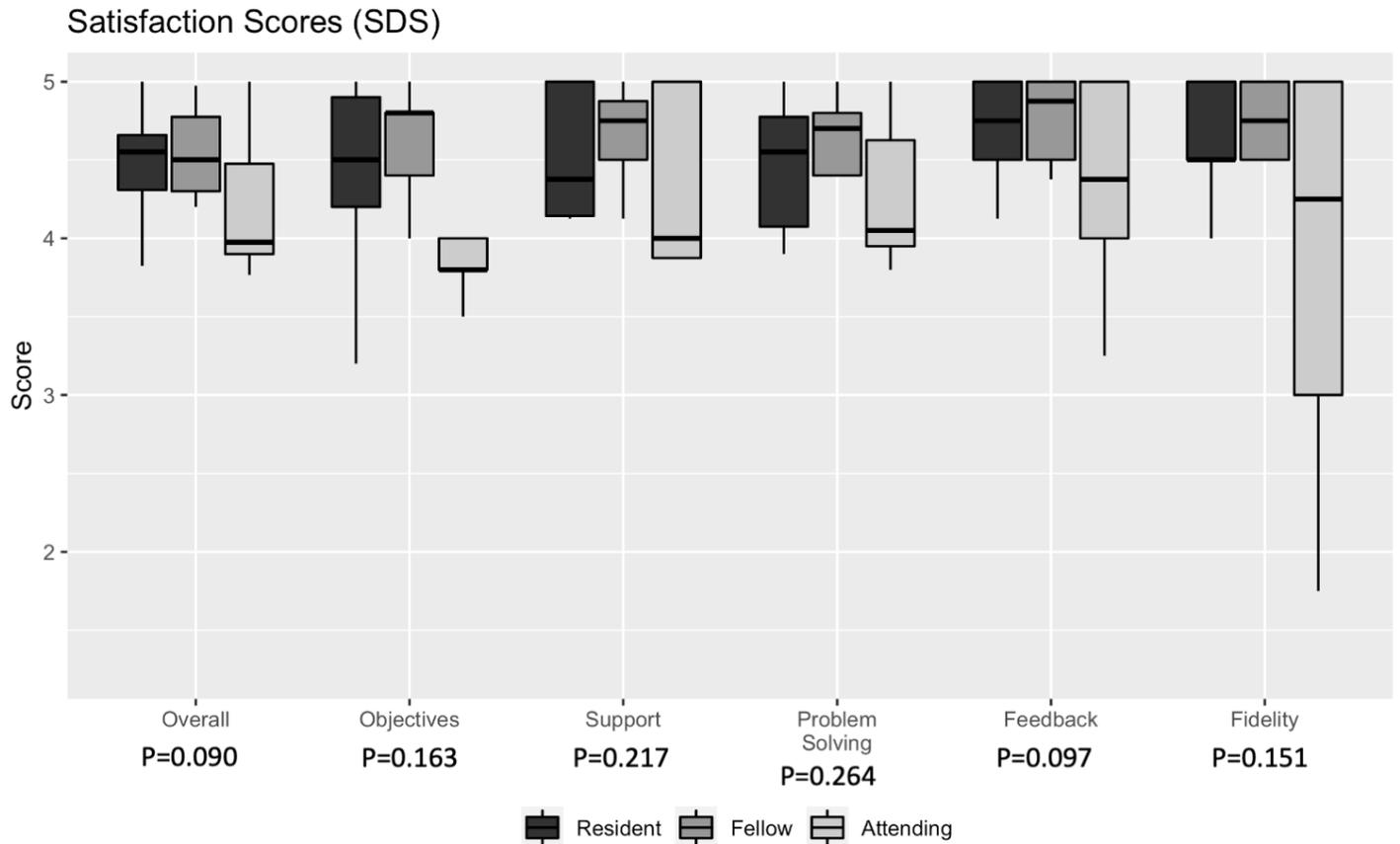
Figures *continued*

Figure 3. *Box plots of SDS scores. The overall SDS score, as well as the section scores, are included. Median scores are grouped by training level.*

Table

Table 1. Demographics

Characteristic	Total, n	Attendings, n	Fellows, n	Residents, n
Age (years ± SD)	36.8 ± 10.1	45.6 ± 13.0	35.1 ± 3.4	29.5 ± 1.7
Sex				
Male	17	5	4	8
Female	10	4	5	1
Race[a]				
American Indian or Alaskan Native	0	0	0	0
Asian	13	5	3	5
Black or African American	1	0	0	1
Native Hawaiian or Other Pacific	0	0	0	0
White	13	4	6	3
Ethnicity				
Hispanic or Latino	2	0	1	1
Not Hispanic or Latino	25	9	8	8
Training Level				
PGY1	0	0	0	0
PGY2	0	0	0	0
PGY3	8	0	0	8
PGY4	1	0	0	1
PGY5	4	0	4	0
PGY6+	5	0	5	0
Instructor	0	0	0	0
Assistant professor	5	5	0	0
Associate professor	1	1	0	0
Professor	3	3	0	0
Previous Exposure to AR				
None	11	3	5	3
1-2 Times	11	4	3	4
2-5 Times	5	2	1	2

Table *continued*

Level of Resuscitation Certification[a]				
Basic Life Support (BLS)	15	5	5	5
Advanced Cardiac Life Support (ACLS)	23	6	8	9
Pediatric Advanced Life Support (PALS)	17	7	9	1
Neonatal Resuscitation Program (NRP)	6	1	4	1
Number Times Initiated Resuscitative Efforts on a Person				
0	1	0	1	0
1-2	6	1	1	4
3-5	6	1	2	3
5-10	4	2	1	1
>10	10	5	4	1
Number Times Initiated Resuscitative Efforts on a Mannequin				
0	1	0	1	0
1-2	3	1	1	1
3-5	5	0	2	3
5-10	7	3	1	3
>10	11	5	4	2
Previously Received Training on Effective Communication Skills During Resuscitation				
Yes	25	8	8	9
No	2	1	1	0
Previously Worked as Frontline Healthcare Worker With Direct Contact to Patients Who Were Critically Ill and In Need of Resuscitation				
Yes	23	7	8	8
1-6 months working in this context	1	0	0	1
7-11 months working in this context	0	0	0	0
1-2 years working in this context	4	0	0	4
>3 years working in this context	18	7	8	3
No	4	2	1	1

Abbreviations: AR, augmented reality; PGY, postgraduate year.

[a] Multiple answers were allowed.

Supplemental Online Material

Appendix A. A-PALS Checklists for Hyperkalemia

1. **General Assessment**

	Performed
Review case including anesthetic record, PMH, labs	Yes ☐ No ☐
Check vitals	Yes ☐ No ☐
Check pulse	Yes ☐ No ☐
Assessed and verbalized anesthesia machine settings	Yes ☐ No ☐
Performed physical exam	Yes ☐ No ☐
Performed timeout	Yes ☐ No ☐

2. **Crisis Specific**

	Performed
Announce problem	Yes ☐ No ☐
Called for help	Yes ☐ No ☐
Call for code cart/defibrillator	Yes ☐ No ☐
Decrease volatile anesthetic	Yes ☐ No ☐
Increase Fi0 2	Yes ☐ No ☐
Review recent medications or fluids administered	Yes ☐ No ☐
Send labs (ABG, VBG, CBC, iSTAT, etc)	Yes ☐ No ☐

Supplemental Online Material *continued*

3.   Hyperkalemia Specific Management

	Performed	
Assessed for hyperkalemia	Yes ☐	No ☐
Started IV fluid bolus	Yes ☐	No ☐
Gave calcium chorlide 10·20 mg/kg or calcium gluconate 30 MG/ KG	Yes ☐	No ☐
Gave excessive dose of calcium	Yes ☐	No ☐
Gave inadequate dose of calcium	Yes ☐	No ☐
Gave Bicarb l-2 meq/kg	Yes ☐	No ☐
Gave excessive dose of Bicarb (>2 meq/kg)	Yes ☐	No ☐
Gave inadequate dose of Bicarb (<l meq/kg)	Yes ☐	No ☐
Gave insulin 0.1 unit/kg	Yes ☐	No ☐
Gave excessive dose of insulin (>0.5 units/kg)	Yes ☐	No ☐
Administered Dextrose 0.5 gm/kg	Yes ☐	No ☐
Administered excessive Dextrose	Yes ☐	No ☐
Administered inadequate Dextrose	Yes ☐	No ☐
Administered Albuterol (5-10 puffs}	Yes ☐	No ☐
Considered Furosemide 0.5-1 mg/kg	Yes ☐	No ☐
Gave excessive dose of Furosemide	Yes ☐	No ☐
Gave inadequate dose of Furosemide	Yes ☐	No ☐
Mentions or Stops Potassium Containing Fluids	Yes ☐	No ☐

4.   Errors

	Performed	
Gave insulin without glucose	Yes ☐	No ☐
Other	Yes ☐	No ☐

Supplemental Online Material *continued*

Appendix B. Simulation Design Scale

In order to measure if the best simulation design elements were implemented in your simulation, please complete the survey below as you perceive it. There are no right or wrong answers, only your perceived amount of agreement or disagreement. Please use the following code to answer the questions

The following rating system was used when assessing the simulation design elements:

1. Strongly Disagree with the statement

2. Disagree with the statement

3. Undecided - you neither agree nor disagree with the statement

4. Agree with the statement

5. Strongly Agree with the statement

NA - Not Applicable; the statement does not pertain to the simulation activity performed

1. Objectives and Information

1. There was enough information provided at the beginning of the simulation to provide direction and encouragement.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
2. I clearly understood the purpose and objectives of the simulation	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
3. The simulation provided enough information in a clear matter for me to problem-solve the situation	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
4. There was enough information provided to me during the simulation	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
5. The cues were appropriate and geared to promote my understanding	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA

2. Support

6. Support was offered in a timely manner.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
7. My need for help was recognized.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
8. I felt supported by the teacher's assistance during the simulation.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
9. I was supported in the learning process.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA

3. Problem Solving

10. Independent problem-solving was facilitated.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
11. I was encouraged to explore all possibilities of the simulation.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
12. The simulation was designed for my specific level of knowledge and skills	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
13. The simulation allowed me the opportunity to prioritize *provider* assessments and care.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA
14. The simulation provided me an opportunity to goal set for my patient.	○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ NA

Supplemental Online Material *continued*

4. Feedback/Guided Reflections

15. Feedback provided was constructive.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA
16. Feedback was provided in a timely manner.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA
17. The simulation allowed me to analyze my own behavior and actions.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA
18. There was an opportunity after the simulation to obtain guidance/feedback from the teacher in order to build knowledge to another level.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA

5. Fidelity (Realism)

19. The scenario resembled a real-life situation.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA
20. Real life factors, situations, and variables were built into the simulation scenario.	◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ NA