HHS Public Access
Author manuscript
Disaster Med Public Health Prep. Author manuscript; available in PMC 2020 May 17.

Published in final edited form as:
Disaster Med Public Health Prep. 2019 April ; 13(2): 301–308. doi:10.1017/dmp.2018.58.

Effects of Virtual Reality Simulation on Worker Emergency Evacuation of Neonates

Sharon Farra, RN, PhD,
Wright State University, Dayton, Ohio

Eric Hodgson, PhD,
Miami University, Miami, Florida

Elaine T. Miller, PhD, RN,
University of Cincinnati, Cincinnati, Ohio

Nathan Timm, MD,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Whittney Brady, DNP, RN,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Matt Gneuhs, MS, CHEP,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Jun Ying, PhD,
University of Cincinnati, Cincinnati, Ohio

Jackie Hausfeld, MSN, RN,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Emily Cosgrove, RN, BSN,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Ashley Simon, RN, MS,
Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio

Michael Bottomley, MS
Wright State University, Dayton, Ohio

Abstract

Objective—This study examined differences in learning outcomes among newborn intensive care unit (NICU) workers who underwent virtual reality simulation (VRS) emergency evacuation training versus those who received web-based clinical updates (CU). Learning outcomes included a) knowledge gained, b) confidence with evacuation, and c) performance in a live evacuation exercise.

Correspondence and reprint requests to Sharon Farra, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH (Sharon.farra@wright.edu).

Methods—A longitudinal, mixed-method, quasi-experimental design was implemented utilizing a sample of NICU workers randomly assigned to VRS training or CUs. Four VRS scenarios were created that augmented neonate evacuation training materials. Learning was measured using cognitive assessments, self-efficacy questionnaire (baseline, 0, 4, 8, 12 months), and performance in a live drill (baseline, 12 months). Data were collected following training and analyzed using mixed model analysis. Focus groups captured VRS participant experiences.

Results—The VRS and CU groups did not statistically differ based upon the scores on the Cognitive Assessment or perceived self-efficacy. The virtual reality group performance in the live exercise was statistically ($P < .0001$) and clinically (effect size of 1.71) better than that of the CU group.

Conclusions—Training using VRS is effective in promoting positive performance outcomes and should be included as a method for disaster training. VRS can allow an organization to train, test, and identify gaps in current emergency operation plans. In the unique case of disasters, which are low-volume and high-risk events, the participant can have access to an environment without endangering themselves or clients.

Keywords

virtual reality; evacuation; disaster

In 2017 there were 132 federally declared major disasters in the United States that resulted in widespread community destruction.[1] In particular, Hurricane Irma forced movement of patients from 35 hospitals, while Hurricane Harvey forced the evacuation of 10 critical infants from Driscoll Children's Hospital in Texas.[2,3] In 2012, neonatal nurses at the New York University Hospital vertically evacuated 19 critically ill infants during Hurricane Sandy.[4] Movement of these fragile newborns was described as both difficult and frightening. As a result of these and other emergency situations, health care workers must adopt an all-hazards approach to training for emergency situations and disasters.[5,6]

The Joint Commission[5] provides management standards for health care organizations to implement to provide patient care through the entire disaster cycle (mitigation, preparedness, response, and recovery). The scope of the preparedness efforts encompasses external (eg, natural disasters; nuclear, chemical, or biological warfare) or internal (eg, fire, water leak, criminal action, explosion) disasters.[5] These standards evolved from studies of significant disasters over the last 5 years and stipulate regular testing of a health care organization's emergency operations plan through realistic and relevant exercise scenarios. The needs of special populations, including elderly, pediatric, and obstetric patients, are specifically highlighted by the Joint Commission.[5]

Emerging evidence suggests that virtual reality simulation (VRS) is a viable method of increasing both learning and retention in emergency preparedness and evacuation training. Virtual reality is immersive and allows participants to partake in training in ways that are not possible with classroom or web-based methods.[7] virtual reality may be more cost-effective and accessible than large-scale real-life exercises. The use of virtual reality is increasing rapidly and research that compares virtual environments with traditional disaster training methods is needed, including exploration of realism, cost/benefits, and translation to disaster preparation and response. VRS has demonstrated effectiveness in many areas of health care, including disaster training.[7] Farra and Miller[8] conducted an integrative review of the literature related to VRS and disaster training and found that several studies demonstrate equal or better learning outcomes with VRS than with traditional classroom teaching. They further found that participants described their experiences as realistic and felt immersed in VRS scenarios.

Virtual simulation has been used in evacuation training to model crowd response in multiple studies.[9‘12] In these studies, the effectiveness of the simulation was not tested, but the simulation was used as the environment to measure other variables. In a virtual evacuation environment, Bode, Wagoum, and Coling[9] explored decision-making with 450 participants. Findings described participant exit choices based on differing time’ dependent information (e.g., queue length and speed at exits). In addition, Andrée, Nilsson, and Eriksson[10] developed a virtual evacuation simulation to assess participant responses to exit choice from a high-rise building. In other evacuation research, VRS was tested as a training method. Serious gaming was shown to be effective in training participants in evacuation from a high-rise building.[11] Garrett and MacMahon[12] reported learning transfer of miners using VRS to train for mine evacuation.

OBJECTIVE

The timely and effective evacuation of neonates during emergencies and disasters is an urgent priority highlighted by the Department of Homeland Security, the Department of Health and Human Services, and recent community experience.[6] The Joint Commission requires biannual activations of a hospital’s emergency operations either in real or simulated events. Patient movement due to mass casualty or evacuation is essential to test the limits to these plans.[2] There is a paucity of research and evidence related to adequate preparation of health care workers in evacuation of neonates. The purpose of this project was to develop and study the effectiveness of a program of VRS training to foster health care workers’ emergency preparedness and response, primarily focusing on the evacuation of neonates from intensive care units. Within this study, we examined differences in learning outcomes and retention among health care workers who underwent VRS evacuation training for neonates and those who reviewed clinical updates which contained the same content as the VRS. Supporting web modules developed for this project provided the baseline training for both groups. The following are the specific aims discussed in this article:

Specific Aim 1

Evaluate the impact of a program of VRS on health care providers’ reactions (comfort, confidence) and knowledge after participating in VRS.

Specific Aim 2

Evaluate the impact of a program of VRS on health care providers’ behaviors during a hands-on simulation exercise.

METHODS

Participants/Setting

The needed sample size was determined by power analysis performed using the G*Power© software. The power level was set to 0.80, and the a was set to 0.05. For the desired effect size to be detected, a total sample size of N = 62 was required (31 per group). Due to the longevity of the study and the potential of loss of participants (over 12 months), an increase of 20% in sample size was determined to be appropriate. The preferred sample size was N = 74.

The study took place on a NICU unit of a large children's hospital. The original sample consisted of 93 participants including registered nurses, monitor technicians (unlicensed personnel), respiratory therapists, physicians, and advanced practice registered nurses. Attrition occurred during the year-long study with participants completing final assessments at 12 months. The attrition rate was 33%, but stratification over treatment group and role was stable. The sample was predominately white female participants (98%). Registered nurses (64%) comprised the largest professional group in the study. The majority of participants were within the 25–31 year age range. Following Institutional Review Board (IRB) review, the consent of the participants was obtained prior to admission to the study.

Workers were enrolled in the study if they were over 18 years of age and worked in the NICU at the children's hospital. The population pool from which subjects were recruited was approximately 330 total employees of the NICU. Enrollment in the study was voluntary, and participants were compensated for their time. NICU nursing and medical leadership did not know if employees chose to participate in the study and had no access to any individual's training or assessment scores. Following completion of the pre-assessments (cognitive assessment and self-efficacy assessment), participants were randomly divided into VRS or clinical update treatment groups with stratified randomization by pre-assessment score results and job classification.

Study Design

The longitudinal experiment used mixed methods (quantitative dominant) to study 2 groups with repeated measures taken at each time step (Table 1). This article presents the quantitative findings. Independent variables included the repeated VRS and the clinical updates administered at 0, 4, 8, and 12 months. Dependent variables were scores on the Emergency Preparedness Information Questionnaire (EPIQ) and cognitive assessment (multiple-choice test) at (pre/post[0], 4, 8, and 12 months). Performance was assessed with live evacuation exercises before treatment and 12 months after treatment. Following the simulation program and final evacuation exercise, VRS participants' qualitative experiences were collected using focus groups.

Measures

EPIQ—The EPIQ measures nurses' self-reported familiarity with emergency preparedness competency dimensions. Wisniewski, Dennik-Champion, and Peltier[13] developed the tool, which was modified by Garbutt, Peltier, and Fitzpatrick.[14] When used previously the EPIQ demonstrated good reliability (R = 0.734). The present study added 12 items specific to neonatal evacuation to the original items. The reliability of the instrument was high with each administration of the modified EPIQ (pre and 0, 4, 8, and 12 months). At 12 months, the assessment of reliability using Cronbach alpha for all EPIQ items was R = 0.978, and for the added 12 items the Cronbach alpha was R = 0.952. The questionnaire was administered through the hospital's secure learning management system.

Exercise Evaluation Tool—Live evacuation exercises using mannequins of newborns were conducted to evaluate evacuation skills. Psychomotor skills were assessed using a rubric developed in collaboration with disaster experts using the Cincinnati Children's Emergency Preparedness and Response Program.[15] Participants vertically evacuated 3 simulated neonates with varying levels of care needs. The researcher-developed tool was used by the raters to evaluate performance. The interrater reliability of the tool was examined using generalizability theory G coefficient (0.984, 95% CI: 0.948–0.9952). The tool also demonstrated high internal consistency with a Cronbach a of (0.845). In the initial exercise, which was conducted prior to any training, scores for the groups that evacuated all 3 neonates ranged from 46 to 83 out of 110 items.[15]

Cognitive Assessment—The pre/post (0 month) assessments question and learning modules were initially developed by the researchers and were based upon the course objectives for the web-based modules. The multiple-choice questions addressed the knowledge/ comprehension (20%) and application (80%) levels of Bloom's Taxonomy.[16] Questions were reviewed and revised for content validity by coinvestigators and external consultants with expertise in disaster training.[15] Revisions were made on the basis of the feedback. The assessment was piloted with a group of hospital staff. After item analysis and revision, all item discrimination scores were positive and above 2. The pilot of the assessment showed good preliminary reliability with a Cronbach alpha of 0.72.[15] Further reliability testing was done with generalizability theory[18] after the final administration of the assessments to examine the variability in item scores to determine how much variability is due to item differences.

Generalizability theory (GT) expands classical reliability theory identifying and estimating the strength of multiple sources of error. GT uses G studies to estimate the magnitude of many potential sources of measurement error. Within GT, the G study isolates and estimates facets of measurement error. The D study uses information from the G study to identify methods to minimize error for a particular purpose.[18] A G-study was undertaken to explore reliability of the cognitive assessment. The G studies for the clinician and monitor technician assessments demonstrated that large variance (69.72%, 61.557) was due to the differences between items. The results of the G-studies were further examined using a D study which resulted in a stability G-score of 0.9528 and 0.835, indicating high reliability.

VARK—The VARK inventory[19,20] was used to assess learning styles of the participants. The VARK measures 4 perceptual preferences: visual (V), aural (A), read/write (R), and kinesthetic (K). Confirmatory factor analysis of the VARK indicated there was support for the validity of the 4 categories of VARK scores.[19] In addition, the reliability coefficients measuring internal consistency for the VARK subscales ranged from .85 to .77 (V, A, R, K) in the investigation.[19] Additional research assessing the internal validity of the VARK learning preferences subscales applied the Rasch model and generated further support of the VARK's appropriateness to measure learning preferences.[20]

Interventions

Web-based Modules—Three modules were developed to advance the training available to staff related to NICU evacuation. Both the treatment group and the control group completed initial web training in neonate evacuation. The web-based modules consisted of the following: (1) overview of hospital policy related to evacuation including types and routes of evacuation, (2) review of neonatal risks during evacuation and neonate preparation and equipment/supplies needed for evacuation, and (3) video instructions for bundling and preparing the neonate for transport including donning the evacuation basket. The modules incorporated best practices of online education, including an unfolding case study and learner interactions.[21]

Virtual Reality Simulations—A series of VRSs (4) (see image Figure 1) were created that augmented the materials presented in the web-based modules. The simulations could be used as either a mouse and keyboard version or using a head-mounted display and a game pad. If participants who chose to use the head-mounted display experienced motion sickness, they were directed to attempt the mouse and keyboard version. If motion sickness continued, they were to stop the simulation. The simulations took approximately 10 minutes to complete and were delivered at 0, 4, 8, and 12 months. The VRS storyboard and content were developed and reviewed by disaster, simulation, and education experts and neonatal clinicians.[22] The objectives, content, and situational challenges contained within the storyboard scaffolded from an easier evacuation scenario (horizontal evacuation of stable infant) to more complex scenarios (evacuation of 3 infants with 2 requiring positive pressure ventilation).

Clinical Updates

The researchers developed clinical updates (CU) using the content from the VRS storyboards. The clinical updates consisted of scenario narratives and embedded questions with included answers. They received visual prompts along with text. The clinical updates thus delivered the same content as the VRS, but in a more traditional medium. Delivery of the clinical updates occurred by computer through the hospital's learning management system at 0, 4, 8, and 12 months (when the VRS group was completing simulations). Participants acknowledged receipt and review of the materials via the learning management system using the assessment tool. These updates contained information that reinforces the content of the initial web-based modules completed by both groups.

Live Disaster Exercise

Using the Department of Homeland Security Exercise and Evaluation Program[23] as a guide, the live exercises researchers worked with NICU and disaster experts to develop scenarios. The scenario included the evacuation of 3 simulated neonates of varying degrees of acuity (stable, ventilator, and oscillator). Implementation of the care of each of the simulated neonates and overall evaluation of the evacuation procedures (correct route, appropriate equipment and communication) were evaluated with the exercise rubric. Trained observers from the hospital staff evaluated participant performance. To assess interrater reliability, each neonate had 2 evaluators assigned to monitor participant actions and 2 individuals monitored overall actions. Job class and treatment group were the basis for participant stratification into the teams, so each team was composed of a mix of job classifications who were all in the same treatment group. Team size (3–6 participants) varied as a result of illness, scheduling, and patient needs. Each team received a score for evacuation, which was a percentage of correct items on the rubrics. Smaller teams (3–4 participants) moved 2 neonates, and full teams (5–6 participants) moved all 3 simulated neonates. Sixteen groups participated in the final exercise, with each scenario lasting approximately 15 minutes. Participants received a prebrief, in which the exercise overview was presented, and a post debriefing, in which feedback was given related to performance.

RESULTS

Specific Aim 1

Evaluate the impact of a program of VRS on health care providers' (1) reactions (comfort, confidence) and (2) knowledge after participating in VRS— Participant confidences represented by the EPIQ[13,14] scores were evaluated using multilevel linear modeling with repeated measures (the assumptions were met). Numerical dependent variables were overall scores and scores of the 12 added items that were specific to neonate evacuation. Mixed effect models showed significant changes of overall emergency preparedness (score on all EPIQ items) ($F_{2,130}$ = 60.64, $P < .0001$) and specific neonatal evacuation items within subjects over time respectively (F_{319}7 = 111.0, $P < .0001$), which indicates that the training, in general, was effective in increasing participants' knowledge. However, between subjects (or between treatment and control groups) differences were not statistically significant on either all EPIQ items ($F_{1\ 130}$ = 1.43, P = .02341) or the 12 items specific to neonatal evacuation ($F_{1,19}$7 = 1.93, P = .1664), which indicates that there were no meaningful differences between the VRS and CU date groups on the EPIQ. The scores appear in Tables 2 and 3.

To evaluate the participant's knowledge of neonate evacuation, the scores on the cognitive assessments were examined using multilinear modeling with repeated measures. Because the number of items was increased for the 8- and 12-month testing periods to improve the assessment reliability, percentage of correct answers was used for final analyses. Both within and between subject changes were examined using mixed effect modeling. As one might expect, scores of participants were statistically different between testing points ($F_{3,19}$5 = 49.79, $P < .0001$), which shows improved knowledge of evacuation procedures as training progressed. There was improvement in scores with each iteration of the intervention with the exception of the post scores in the clinical update group, which were the highest scores overall for the group. The change from the pretreatment scores to the final scores indicates significant improvement in the understanding of concepts of neonatal evacuation. There were no statistically significant differences between participants trained with VRS or clinical updates ($F_{1\ 19}$5 = 0.05, P = .8193). There were also no significant differences detected between groups at any individual time points. Table 4 presents the combined cognitive assessment results with analysis between time points. Additional analysis of the cognitive test considered the effects of role and age. As before, neither age nor role had a statistically significant effect or interaction with training group (all Fs > 0.09).

The student's preferred method of learning was measured using the VARK (visual, audio, reading, and kinetic) learning style assessment. Relationships between the VARK and the overall response variables of EPIQ, cognitive assessment, and exercise scores were examined using linear regression. There was no issue with multicollinearity found within the variables (EPIQ P value range .1277- .711, cognitive assessment P value range .1290-.7434, exercise score P value range .4559.9180). The results of this analysis indicated that there was not sufficient evidence of a linear relationship between any of the response variables and the style of learning. Further analysis of the VARK variables included examination of the relationship of method of learning (VRS, CU) to response variables. Again, the evidence was insufficient to establish a linear relationship between the response variables (EPIQ P values ranged from .1277-.7111; cognitive assessment P values were .275-.803; exercise score P values were .0908.9040). Learning style did not have an apparent effect on performance overall or between the treatment groups.

Specific Aim 2

Evaluate the impact of a program of VRS on health care providers' behaviors during a hands-on simulation exercise—Using linear modeling to determine differences within subjects and between the treatment groups the researchers examined exercise performance scores and time to completion. Both the VRS group and the clinical update group improved significantly from the initial exercise to the second exercise. There was no significant difference in the average time for evacuation in the final exercise between the treatment groups (see Table 5). Both groups had statistically significant improvement in scores on evacuation performance from the first exercise to the final exercise. In addition, the 2 groups differed significantly t (63) = 7.13, $P < .0001$) on performance scores during the final exercise with the VRS group having a better performance (mean scores of VRS: 86.46, SE ±1.5; CU: 71.08 SE ±1.5). Estimation of the effect size is 1.71

DISCUSSION

The results demonstrate a mixed but overall positive result for the effect of VRS on NICU evacuation training. There were no differences between the 2 groups (VRS, CU) based upon measures of self-efficacy (EPIQ) or general knowledge (cognitive assessment). However, actual performance in a live evacuation drill (as measured by the rubric) was both statistically and clinically different between the groups, with the VRS group having notably better performance. The VRS group also evacuated the babies slightly faster and more efficiently than the clinical update group, although this latter effect did not reach statistical significance and may need to be confirmed with further research.

It is unclear at this point why there were such stark differences between the assessment measures and the results of the live drill. It is possible that the cognitive assessment was not able to capture some aspects of neonate evacuation knowledge. Assessing which specific items CU participants tended to miss while VRS participants tended to get correct would help to improve the cognitive assessment in the future. With regards to self-efficacy measures, it seems likely that training in general leads to equivalently high levels of self-confidence in one’s knowledge, regardless of the training medium. In this case, it would not be surprising for EPIQ scores to be equivalent between groups. These issues will need to be addressed in future work. Viewing the data as a whole, however, it is apparent that the VRS group did no worse than traditionally-trained participants on any given measure, and showed a distinct performance advantage on the most direct behavioral measure—performing in a live exercise. As these drills are designed to simulate an actual disaster, one would presume this performance advantage for the VRS group would also be apparent should they ever be called on to use their training. This bodes well for the use of VRS as a training medium in health care generally and for disaster preparedness specifically.

Aside from the advantage for VRS training, the present study also indicated that training in general was effective for all participants. There were improvements on all 3 outcome measures (EPIQ, cognitive assessment, and exercise scores) for subjects in general. Variables of age and role did not reveal any significant differences or interactions between the groups. The examination of learning style measured by the VARK scores likewise indicate that learning style preferences did not have any linear relationship with scores on other measures. Learning style did not appear to affect performance in the exercise and was not a predictor of performance on the assessments. Given that the learning experience for both groups combined multiple learning strategies including auditory, reading, and kinetic modalities, it is not surprising that one single learning style did not appear to affect the EPIQ, cognition, or exercise scores.

Limitations

This study consisted of a convenience sample with the subjects randomly assigned to either the control or the intervention group. The sample was predominately white female participants. The study was conducted at a single site, so findings may not be generalizable to other settings. There was 33% attrition rate from the study, but the final number of subjects met the projected desired sample size. Another threat to internal validity was testing, but this aspect was addressed by having multiple forms of the instrument study questions. One participant developed motion sickness with either mouse and keyboard or oculus VRS and transferred to the clinical update group. The results of the exercise were a team score versus individual scores there may be bias in giving credit to participants who may not have performed well individually. It should also be noted that the NICU ward experienced 2 real partial evacuations during the year-long study period (between the 4 and 8 month training periods), which were unrelated to the study but helped underscore the need for such training. These were unforeseen events that gave on- duty staff extra evacuation experience, but the effects on this study should be equally distributed between groups.

CONCLUSIONS

The interventions provided to both groups resulted in learning. The learning outcomes from the modules, which were reinforced with short boosters from either the VRS or the CU, resulted in retention of content across the one-year span of the study. The VRS demonstrated impact at level 3 of Kirkpatrick's model[24] through application in the simulated clinical environment with VRS participants demonstrating statistical and clinical improvement in performance during the live exercise. The demographic characteristics and learning styles of individuals did not show a relationship to the response variables and the type of training experienced.

Implications

The results of the study add to the body of knowledge related to VRS (mouse and keyboard or head-mounted display) as a disaster training method. The use of VRS enhanced performance. Current evidence supports the use of VRS as a training method, but only a few studies have used experimental designs.[25–30] The reviewed literature related to VRS used in overall evacuation training is predominately focused on model crowd response. This research supports the findings of Ribeiro, Almeida, and Rossetti et al.[11] and Garrett and MacMahon,[12] who effectively used VRS to train participants in evacuation. This study demonstrates the importance of considering more than one measure of outcome performance. The study demonstrated no difference in perceived self-efficacy (measured by EPIQ) or cognitive knowledge (measured by cognitive assessment), but demonstrated a large effect on performance (measured by live exercise performance rubric). If performance had not been measured, the treatment effect on learning would not have been found.

Training using VRS is effective in promoting positive performance outcomes and should be included as a method for disaster training. VRS can allow an organization to train, test, and identify gaps in current emergency operation plans. In the unique case of disasters, which are low-volume and high-risk events, the participant can have access to an environment without endangering themselves or clients. The VRS is perceived as very realistic, which participants identify as an important component of training. Advantages of VRS over traditional types of training include the ability of participants to participate asynchronously from any computer that has the program available. VRS can be constructed to give immediate feedback, provide instruction, and track evaluation information for both the participant and the organization. The participant may access the program as many times as they would like to reinforce learning or improve performance.

Acknowledgments

We thank the Agency for Health Research and Quality (AHRQ) for their support of our project. AHRQ Grant No.: R18 HS23149

REFERENCES

1. Federal Emergency Management Agency. Disaster Declarations by Year. http://www.fema.gov/disasters/grid/year. Published 2017 Accessed December 1, 2017.
2. Blau M Irma forces at least 35 hospitals to evacuate. STAT News. https://www.statnews.com/2017/09/09/irma-hospital-evacuations-rundown/. Published September 9, 2017 Accessed December 1, 2017.
3. Seiger T Critically ill babies evacuated from Texas coast ahead of Harvey. Cox Media Group, National Content Group. http://www.ajc.com/weather/hurricanes/critically-ill-babies-evacuated-from-texas-coast-ahead-harvey/sdqR2GR920QngkRCmH8MmL/. Accessed December 1, 2017.

4. Espiritu M, Patil U, Cruz H, et al. Evacuation of a neonatal intensive care unit in a disaster: lessons from Hurricane Sandy. Pediatrics. 2014;134 (6):e1663–e1669.

5. The Joint Commission. Emergency management standards: supporting collaboration and planning. www.jointcommission.org/assets/1/6/EM_Stds_Collaboration_2016.pdf. Published 2016 Accessed December 1, 2017.

6. Department of Homeland Security. Homeland security grant program supplemental resource: children in disasters guidance. 2012 https://www.fema.gov/pdf/government/grant/2012/fy12_hsgp_children.pdf. Accessed August 3, 2018.

7. Hsu EB, Li Y, Bayram JD, et al. State of virtual reality based disaster preparedness and response training. PLoS Curr. 2013 4 24:5.

8. Farra SL, Miller E. Virtual reality and disaster training: an integrative review of the literature. J Nurs Educ Pract. 2013;3(3):655–671.

9. Bode NWF, Wagoum KAU, Codling EA. Information use by humans during dynamic route choice in virtual crowd evacuations. R Soc Open Sci. 2015;2(1):140410. doi: 10.1098/rsos.14041 [PubMed: 26064589]

10. Andrée K, Nilsson D, Eriksson J. Evacuation experiments in a virtual reality high-rise building: exit choice and waiting time for evacuation elevators. Fire Mater. 2016;40(4):554–567.

11. Ribeiro J, Almeida J, Rossetti R. Using serious games to train evacuation behaviour. Paper presented at: 7th Iberian Conference on Information Systems and Technologies; June 20–21, 2012; Madrid, Spain.

12. Garrett M, McMahon M. Indirect measures of learning transfer between real and virtual environments. Australas J Educ Technol. 2013;29(6):806–822.

13. Wisniewski R, Dennik-Champion G, Peltier JW. Emergency preparedness competencies: assessing nurses' educational needs. J Nurs Adm. 2004;34:475–480. [PubMed: 15577671]

14. Garbutt SJ, Peltier JW, Fitspatrick JI. Evaluation of an instrument to measure nurses' familiarity with emergency preparedness. Mil Med. 2008;173(11):1073–1077. [PubMed: 19055181]

15. Farra S, Miller ET, Gneuhs M, et al. Evacuation performance evaluation tool. Am J Disaster Med. 2016;11(2):131–136. doi: 10.5055/ajdm.2016.0232 [PubMed: 28102534]

16. Bloom BS, Krathwohl DR. Taxonomy of Educational Objectives: The Classification of Educational Goals, by a Committee of College and University Examiners. Handbook 1 : Cognitive Domain New York: Longmans; 1956.

17. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple- choice item-writing guidelines for classroom assessment. Appl Meas Educ. 2002;15(3):309–334.

18. Shavelson RJ, Webb N. Generalizability Theory: A primer. Los Angeles: Sage Publications; 1991.

19. Leite WL, Svinicki M, Shi Y. Attempted validation of the scores of the VARK: learning styles inventory with multitrait-multimethod confirmatory factor analysis models. Educ Psychol Meas. 2010;70:323–339.

20. Fitkov-Norris ED, Yeghiazarian A. Validation of VARK learning modalities questionnaire using Rasch analysis. J Phys Conf Ser. 2015;588:012048.

21. Interactive Petty J., technology-enhanced self-regulated learning tools in healthcare education: a literature review. Nurse Educ Today. 2013;33:53–59. [PubMed: 22818225]

22. Farra S, Miller ET, Hodgson E, et al. Storyboard development for virtual reality simulation. Clin Simul Nurs. 2016;12(9):392–396. doi: 10.1016/j.ecns.2016.04.002

23. Department of Homeland Security: Homeland Security Exercise and Evaluation Program. http://www.fema.gov/media-library-data/20130726-1914-25045-8890/hseep_apr13_pdf. Accessed December 1, 2017.

24. Kirkpatrick D Techniques for evaluating training programs: revisiting Kirkpatrick's four-level model. Train Dev. 1996;50(1):54–59.

25. Dubovsky SL, Antonius D, Ellis D, et al. A preliminary study of a novel emergency department nursing triage simulation for research applications. BMC Res Notes. 2017;10:101–112. doi: 10.1186/s13104-016-2337-3 [PubMed: 28209187]

26. Mohan D, Rosengart MR, Fischoff B, et al. Testing a videogame intervention to recalibrate physician heuristics in trauma triage: study protocol for a randomized controlled trial. BMC Emerg Med. 2016; (44):1–10. [PubMed: 26728878]

27. Mohan D, Baruch F, Farris C, et al. Validating a vignette-based instrument to study physician decision making in trauma triage. Med Decis Making. 2014 2;34(2):242–252. [PubMed: 24125789]

28. Ingrassia PL, Ragazzoni L, Carenzo L. Virtual reality and live simulation: a comparison between two simulation tools for assessing mass casualty triage skills. Eur J Emerg Med. 2015;22(2):121–127. [PubMed: 24841770]

29. Cohen D, Sevdalis N, Patel V, et al. Tactical and operational response to major incidents: feasibility and reliability of skills assessment using novel virtual environments. Resuscitation. 2013;84(7):992–998. [PubMed: 23357703]

30. Pucher PH, Batrick N, Taylor D, et al. Virtual-world hospital simulation for real-world disaster response: design and validation of a virtual reality simulator for mass casualty incident management. J Trauma Acute Care Surg. 2014;77(2):315–321. [PubMed: 25058260]

FIGURE 1.

Virtual Reality Simulation Screen Shot.

TABLE 1

Timeline

Activity	Pre Proc.	Treatment	0 Month	4 Month	8 Month	12 Month
EPIQ[a]	X		X	X	X	X
Cognitive	X		X	X	X	X
assessment				(12 items)		
Live exercise	X					X
3 learning modules		X				
Clinical update		X		X	X	X
OR VRS						
Focus group						X (VRS)

[a]Abbreviations: EPIQ, Emergency Preparedness Information Questionnaire; VRS, virtual reality simulation.

TABLE 2

Emergency Preparedness Information Questionnaire[a]: All Items

	Mean ± SE			*P* Value		
Training Status	**Pre**	**Post (0 Month)**	**12 Month**	**Pre/Post**	**Post/12 Month**	**Pre/12 Month**
CU[b]	216.5 ±6.3	180.1 ±6.9	170.64 ±7.4	<.0001	.8536	<.0001
VRS	218.88±6.3	168.6±6.8	153.4 ±7.6	<.0001	.4442	<.0001
P value	.9998	.8402	.5787			

[a]The scale is reverse scored: lower scores indicate higher confidence.

[b]Abbreviations: CU, clinical updates; VRS, virtual reality simulation.

TABLE 3

Emergency Preparedness Information Questionnaire[a]: 12 Items

	Mean ± SE			*P* Value		
Training Status	**Pre**	**Post (0 Month)**	**12 Month**	**Pre/Post**	**Post/12 Month**	**Pre/12 Month**
CU[b]	38 ± 1.2	24 ± 1.3	22.55 ±1.3	<.0001	<.0001	.9903
VRS	39.68 ±1.2	19.18± 1.3	20.88 ±1.4	<.0001	<.0001	.9806
P value	.9793	.1815	.9893			

[a]The scale is reverse scored: lower scores indicate higher confidence.

[b]Abbreviations: CU, clinical updates; VRS, virtual reality simulation.

TABLE 4

Cognitive Assessment: Combined Results Clinician & Monitor Technician

	Mean ± SE				*P* Value[a]			
Training Status	**Pre**	**Post (0 Month)**	**8 Month**	**12 Month**	**Pre/Post**	**Post/8 Month**	**8/12 Month**	**Pre/12 Month**
CU[b]	58.75 ±1.6	72.76 ±1.8	68.48 ±1.8	70.74 ±1.9	<.0001	.7716	.9992	<.0001
VRS	55.59 ±1.6	68.48 ±1.8	71.91 ± 1.8	74.19 ±1.9	<.0001	.7554	.973	<.0001
P value	.8681	.6856	.9802	.9081				

[a]Statistical significance, $P < .05$.

[b]Abbreviations: CU, clinical updates; VRS, virtual reality simulation.

TABLE 5

Exercise Time and Score

	Treatment Group	Score	Time, min
Mean±SE	I. Pretraining	63.27 ±2.2	19 ± 1.4
	II. CU	71.08± 1.5	14 ± 1
	III. VRS	86.46 ±1.5	13 ±1
Statistical differences[a]	I vs II	$t(63) = 2.94$, $P=.0126$	$t(63) = 2.47$, $P=.0425$
	I vs III	$t(63) = 8.76$, $P<.0001$	$t(63) = 3.14$, $P=.0072$
	II vs III	$t(63)= 7.13$, $P<.0001$	$t(63) = 0.8$, $P=.7019$

[a] Statistical significance, $P<.05$.