



Using Virtual Simulations to Assess Situational Awareness and Communication in Medical and Nursing Education: A **Technical Feasibility Study**

Steven J. Anbroa, Alison J. Szarkoa, Ramona A. Houmanfara, Amber M. Maraccinib, Laura H. Crosswell^c, Frederick C. Harris^d, Michelle Rebaleati^e, and Luka Starmer^e

^aUniversity of Nevada, Behavior Analysis Program, Reno, Nevada, USA; ^bRenown Health, Reno, Nevada, USA; 'University of Nevada, Reynolds School of Journalism, Reno, Nevada, USA; 'University of Nevada, Department of Computer Science & Engineering, Reno, Nevada, USA; eUniversity of Nevada, @One Digital Media, Reno, Nevada, USA

ABSTRACT

Communication and situational awareness, among other "human factors," are critical skills needed within high-reliability organizations (HROs). HROs are challenged to develop an effective methodology for the systematic assessment of these skills. Virtual reality (VR) simulation technology offers a promising approach to meet this challenge. By utilizing a verbal coding procedure in tandem with eye-tracking metrics, we conducted a technical feasibility study to assess the impact of an interprofessional training (TeamSTEPPS®) on communication accuracy and observing responses among medical and nursing students in a virtual simulation. The results suggest that communication accuracy significantly improved as a result of TeamSTEPPS® training. These findings and changes in situational awareness among individuals are discussed.

KEYWORDS

Virtual reality; simulation; communication; situational awareness; medical education

The science of human behavior relies on empirical research to further our understanding, prediction, and influence of behavioral phenomena. The use of analog environments for research often results in more experimental control than natural environments. The role of simulations is therefore increasingly relevant. A review of simulation-based, empirical research published in behavior analytic journals (Ward & Houmanfar, 2011) identified 46 studies (avg. 2 per year) between 1987-2010. We extended this review, identifying additional 49 simulation-based studies in behavior analytic journals (avg. 7 per year) between 2011–2018. Despite the steadily increasing rate of simulation-based studies within behavior science, none to date have used virtual reality (VR). VR provides an immersive audio/visual experience within a virtually simulated environment. This technology has steadily gained popularity for gaming applications; in research applications, simulations are entirely under the control of the programmer/experimenter. Advances in VR

include the addition of eye-tracking sensors within VR headsets, allowing data collection that was heretofore impossible. VR with eye-tracking allows researchers to record objective measures conceptually linked to behavioral phenomena that lack empirical measurement.

Situational awareness is one such phenomenon, comprised of many individual responses which "attempt to explain human behavior in operating complex systems" (Endsley, 1995a, p. 32). Situational awareness is a critical skill within high-reliability organizations (HROs): organizations whose employees conduct highly technical operations in working conditions with varying levels of risk (Alavosius, Houmanfar, Anbro, Burleigh, & Hebein, 2017). The predominant conceptual model of situational awareness (Endsley, 1995a, 1995b) depicts three "levels" of awareness that build upon one another: Level 1) perception of elements in the current situation, Level 2) comprehension of the current situation, and Level 3) projection of future status. Killingsworth, Miller, and Alavosius (2016) reinterpreted Endsley's model from a behavioral perspective. To summarize: Level 1 is conceptualized as stimulus control, conditional discriminations, and observing responses; Level 2 is conceptualized as tacting stimuli functions/features and the interlocked behavioral chains that follow; Level 3 is conceptualized as predicting responses and performance generalization across contexts. To date, methods for assessing situational awareness rely primarily on subjective measures. However, through advances in VR technology, researchers may begin studying this phenomenon more objectively, with high precision and reliability.

Behavioral measurement and assessment of communication accuracy has a richer history and methodology to draw from. Bijou, Chao, and Ghezzi (1988) introduced a referential verbal coding system that was later used by Johnson, Houmanfar, and Smith (2010) to assess measures of productivity and customer service. This procedure involved the transcription of audio/ video recordings, development of operationally defined coding categories, and scoring of the transcribed segments with respect to the correctly identified coding category. Smith, Houmanfar, and Denny (2012) refined this method, allowing researchers to quantify communicative statements exchanged between pairs for enhanced measurement: frequencies of correct responding and the duration of verbal interactions. Most recently, Maraccini, Houmanfar, Kemmelmeier, Piasecki, and Slonim (2018) used this refined method to assess communication accuracy among medical/nursing student pairs during simulated patient handoffs. Continued use of this assessment method is warranted, as it provides an objective assessment of communication accuracy in team settings.

A current "gold standard" training within the medical industry, designed to improve both situational awareness and communication accuracy, is TeamSTEPPS® (AHRQ, 2008; Epps & Levin, 2015; King et al., 2008; Maraccini et al., 2018; Sheppard, Williams, & Klein, 2013). TeamSTEPPS® was developed in response to the high rate of medical errors leading to "sentinel events": unintended patient death or non-fatal complications resulting from medical intervention (Maraccini et al., 2018). The prevalence of death within the United States due to medical error (Makary & Daniel, 2016) emphasizes the need for a remedy. While TeamSTEPPS® has yielded promising outcome improvements among medical teams (e.g., Sawyer, Laubach, Hudak, Yamamura, & Pocrnich, 2013; Turner, 2012), objective assessment of both situational awareness and communication accuracy is absent. VR may help mitigate this limitation. Therefore, the purpose of the current study was to assess the impact of an interprofessional TeamSTEPPS® training on communication accuracy and observing responses among medical and nursing students in a virtual simulation.

Method

Design, participants, & setting

A pre/post comparative design was used to evaluate changes in communication accuracy during an emergency medical simulation as a function of an interprofessional TeamSTEPPS® training. The study was conducted on a medical school campus, using five standard patient rooms (equipped as fully functioning clinician offices) to run simultaneous VR simulations. One hundred five students (49 third-year medical students and 56 third-year nursing students; 63 females, 42 males) participated in this study. Student participation was built into curriculum requirements by the medical and nursing schools involved in this study. A pre-training VR simulation was followed by an interprofessional TeamSTEPPS® training workshop in a large lecture hall. Approximately 1 week after training, students completed a posttraining VR simulation.

Materials

Five HTC Vive VR headsets were used, one of which was equipped with integrated Tobii eye-tracking sensors. The integration of the Tobii eyetracking hardware is comprised of one eye-tracking sensor per eye built into the HTC Vive headset. These sensors operate using the binocular dark pupil tracking technique (Tobii Pro, n.d.) to detect absolute pupil size, pupil position, gaze direction, gaze origin, and time for each eye fixation. There is an automated 5-point calibration procedure that is conducted to ensure accuracy prior to each participant's use of the eye-tracking hardware, which takes approximately 1-2 min on average.

A 5 min., 42 sec. VR simulation was filmed using a 360° video camera prior to the onset of the study. In this simulation, a life-sized mannequin (SimMom) served as a simulated patient; a SimMom is a medical training device that is capable of fully simulating birth and various complications that researchers control, such as an eclamptic seizure (used in the current simulation). The SimMom's actions, vital readouts, and verbal responses were managed from a control room outside the simulated patient room where filming occurred. An experienced medical technician performed the role of a nurse in the simulation. The role of the physician in the simulation was performed by an experienced OB/GYN medical doctor. The script used during filming was crafted by a second experienced OB/GYN medical doctor, ensuring the adherence to current standards of medical practice for patients experiencing an eclamptic seizure.

Procedure

Pre/post-training medical simulation

Participants entered a simulated patient room and were fitted with an HTC Vive VR headset. Five standard patient rooms were used simultaneously. Research assistants visually calibrated the HTC Vive headsets for each participant, before providing verbal instructions. Participants were told to say "start" aloud when they saw a visual prompt in the headset, beginning the simulation. They were also told to respond aloud to the simulation if and when they see fit. In the simulation itself, students saw a physician and nurse interact with the SimMom (see Figure 1) for 3 min and 27 s, gathering intake



Figure 1. Sample still image from the virtual simulation used in the current study. In this simulation, a nurse (left: Alicia Fong; University of Nevada, Reno School of Medicine) and physician (right: Dr. Earle Oki; Renown Health) treat a simulated patient (center; "SimMom" fullbody birthing simulator) who is experiencing an eclamptic seizure. The observer (participant) is asked by the nurse to provide verbal checkbacks as treatment progresses.

information, and measuring vital signs. At this point in the simulation, the SimMom was programmed to have an eclamptic seizure. The onset of the eclamptic seizure began the crisis event (i.e., the medical emergency) segment of the simulation, which lasted 2 min and 5 s. At the onset of the crisis event, the simulated nurse addressed the medical or nursing student participant viewing the simulation by looking and speaking directly into the camera, and asking the student participant to serve as a recorder of the event by providing verbal checkbacks (i.e., verbal confirmation of procedural steps). Students could verbally respond at any point during the simulation; verbal responses were measured and assessed for accuracy during 10 predetermined "verbal checkback opportunities." Table 1 shows a timeline of simulation events and predetermined verbal checkback opportunities.

Participants in one of the five simulated patient rooms (N = 22) used an HTC Vive headset equipped with eye-tracking sensors as described above; only one of the five HTC Vive headsets was equipped with eye-tracking sensors due to limited availability of hardware at the time of this study. Participants assigned to the eye-tracking headset were chosen at random from the full 105 participants, using a random number generator. The eyetracking calibration process was initiated within the headset's display, following visual calibration by research assistants. The headset provided visual prompts: a series of dots that participants were instructed to look at. This calibration process within the headset allowed the eye-tracking sensors to accurately capture metrics of participants' eye gazes within primary areas of interest (AOIs) during the simulation. AOIs were predetermined to be critical visual elements of the simulation that participants would need to attend to namely, the physician and the nurse (the speakers), during the crisis event.

Interprofessional training

The Chief Nursing Officer of a large local hospital and the Assistant Program Director for a local medical school conducted an interprofessional training session for the medical and nursing student participants. This training included material from the full TeamSTEPPS® curriculum (see AHRQ, 2012), condensed into a 3-h workshop. The workshop included lectures over the patient safety movement and TeamSTEPPS® model/core skills. A discussion of techniques for improving situational awareness and communication accuracy was followed by brief practice sessions. These practice sessions tasked participants with constructing some end-product (e.g., the tallest house of cards possible, a specific Mr. Potato Head configuration) using decreasingly restrictive communication. Three rounds of practice sessions included certain communication rules: 1) no talking; 2) only an appointed team leader may talk; 3) all team members may talk.

Table 1. Timeline of events in the simulation used. Start times and stop times are shown for simulated events and verbal statements from the physician/nurse. Sample verbal checkbacks (participant responses) meeting "correct" criteria are shown; accuracy criteria for "correct" responses are shown in the "Sample Verbal

Checkb	ack" c	Checkback" column, in parentheses. All criteria must be met for a response to be scored as "correct."	
Time	Time	Physician (MD) Statements, Nurse (RN) Statements,	
Start	Stop	& In-Simulation Events	Sample Verbal Checkback
0:00	0:13	(Instruction appears: "Please say START after the countdown", followed by a 3-2-1 visual countdown, then " START ")	"Start." (Used for video recording calibration)
0:13	3:26	(Simulation begins: RN briefs MD on patient vitals; MD asks the simulated patient several intake questions; e.g.,	N/A
		"Have you ever head any heart, liver, or kidney problems in the past?")	
3:27	N/A	("CRISIS EVENT" – Simulated patient seizure begins)	N/A
3:28	3:30	MD: "She's having an eclamptic seizure."	(optional) "Providing verbal checkbacks."
3:31	3:35	RN: "Student, I'm going to need you to be a recorder and provide verbal feedback for us."	
3:38	3:41	MD: "Let's have her in a left lateral decubitus."	1. "(1) Left lateral decubitus."
		(MD & RN move patient into a left lateral decubitus position)	
3:45	3:46	RN: (moves bed rails up into position) "Bed rails up."	2. "(1) Bed rails up."
3:46	3:47	MD: (moves bed rails up into position) "Bed rails up."	
3:48	3:51	(RN alerts charge nurse & anesthesia)	3. "(1) Charge nurse and (2) anesthesia contacted."
		RN: "I've contacted the charge nurse and anesthesia."	
3:59	4:01	(MD places oxygen mask on patient)	4. "(1) Oxygen on, (2) airway protected."
		MD: "Oxygen on, airway is protected."	
4:04	4:06	RN: "Oxygen is at 10 liters."	5. "(1) Oxygen at (2) 10 liters."
4:07	4:09	MD: "Ok, let's confirm she has no contraindications to magnesium sulfate."	6. "(1) No contraindications to magnesium sulfate."
4:10	4:12	RN: (checks patient chart) "She has no contraindications."	
4:17	4:23	MD: "Let's go ahead and give her 6 grams magnesium sulfate IV over 15 minutes, please."	7. "(1) 6 grams (2) magnesium sulfate (3) IV (4) over
4:24	4:28	(RN holds medication up to participant, showing label)	15 minutes."
		RN: "6 grams magnesium sulfate IV over 15 minutes."	
4:33	4:34	(RN begins administering medication; patient seizure ends) MD: "The seizure has stopped."	8. "(1) Seizure has stopped."
5:26	5:30	(MD briefs patient on what happened & answers her questions; e.g., "Is my baby ok?")	9. "(1) Magnesium sulfate (2) on board for (3) 6
		(RN continues administration of medication)	grams running."
		RN: "Magnesium sulfate is on board, 6 grams running."	
5:34	5:36	MD: "Fetal heartrate is returning to baseline."	10. "(1) Fetal heartrate (2) returning to baseline."
N/A	5:42	(Simulation ends)	N/A



Dependent variables

The primary dependent measure was communication accuracy. All participants were audio/video recorded during their observations of the medical simulation. Video recordings were then transcribed and each of the 10 predetermined opportunities for verbal checkbacks was assessed for accuracy, using the specific accuracy criteria shown in Table 1. "Correct" responses were those which met all accuracy criteria; to be considered "correct," verbal responses also had to begin within 2 seconds of the most recent statement made by the physician/nurse. "Erred" responses were those in which at least one of the listed accuracy criteria contained incorrect information. "Missing" responses were those in which a) multiple accuracy criteria are present, and b) at least one of those criteria was absent from the verbal response. "Omitted" responses were recorded if a participant did not say anything during a verbal checkback opportunity.

Hollowell and Lansing (2004) describe the need for temporal accuracy when measuring eye-tracking during fast processes (e.g., a medical crisis). Crosswell, Porter, and Sanders (2018) emphasized the utility of eye-tracking metrics such as eye fixation duration, which was chosen as a secondary dependent measure for those participants using the eye-tracking enabled HTC Vive (N = 22). Eye-tracking data were automatically collected via the Tobii eye-tracking sensors integrated within the HTC Vive headset. Eyetracking sensors captured eye fixations where a participant's pupils were directed at one of the AOIs for at least 250 ms (0.25 s). Raw eye fixation data from the crisis event segment of the simulation were analyzed using Tobii software to determine fixation duration (inspireUX, 2010) of eye fixations on the physician and nurse AOIs.

Results and discussion

Figure 2 shows a stacked bar graph of all 105 participants' coded verbal responses, during both pretest and posttest, at each of the 10 verbal checkback opportunities. Verbal responses shown in this graph have been grouped into "correct" responses (black bar segments) and "erred/missing/ omitted" responses (white bar segments). At each checkback opportunity, the number of participants emitting correct verbal responses increased from pretest to posttest. Trial-by-trial inter-observer agreement was calculated for 33% of both pretest and posttest sessions. Pretest sessions yielded 92% agreement (range, 80-100%) and posttest sessions yielded 91% agreement (range, 70-100%). A paired-samples t-test was conducted to compare communication accuracy before and after training. There was a significant difference between pretest (M = 1.72, SD = 2.49) and posttest (M = 4.14, SD = 3.11) conditions; t = 8.2734, p < .0001. While the number of students

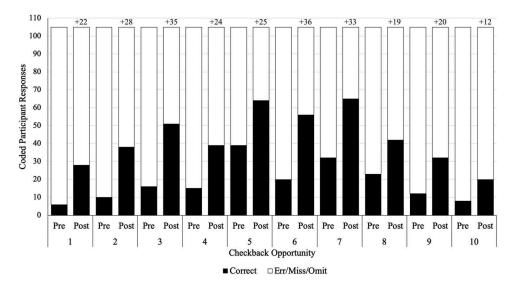


Figure 2. Participants' coded verbal responses across 10 verbal checkback opportunities in pretest & posttest. Black bars indicate the number of participant responses meeting "correct" criteria, white bars indicate the number of participant responses meeting "erred," "missed," or "omitted" criteria. Pretest to posttest differences of correct responses are shown above the posttest bars.

emitting correct verbal responses increased, it should also be noted that there was substantial room for improvement. Out of 10 checkback opportunities, only two in posttest (#5 & 7) resulted in over 50% of participants responding correctly. Clinically, these results are lacking. Medical and nursing students would no doubt benefit from ongoing training of standard communication techniques, practice opportunities in various simulations, and repeated assessment of communication accuracy – individually and in groups.

Figure 3 shows a scatterplot of those 22 participants who used the HTC Vive headset with embedded eye-tracking. The x-axis shows the change in eye fixation duration on the speaker as they are speaking, from pretest to posttest. Data points on the right of the scatterplot indicate those participants looked at the speaker for a longer duration in posttest than pretest. Data points on the left indicate a decrease in time spent looking at the speaker. The y-axis shows the change in each participant's number of correct verbal responses from pretest to posttest. Data points above 0 indicate those participants who emitted more correct responses in posttest than in pretest; data points below 0 indicate those participants who emitted fewer correct responses in posttest than in pretest. Seventeen participants increased their frequency of correct verbal responses in posttest: 5 looked at the speaker longer, 11 looked at the environment longer, 1 did not change. Four participants did not increase their frequency of correct verbal responses in posttest: 2 looked at the speaker longer, 2 looked at the

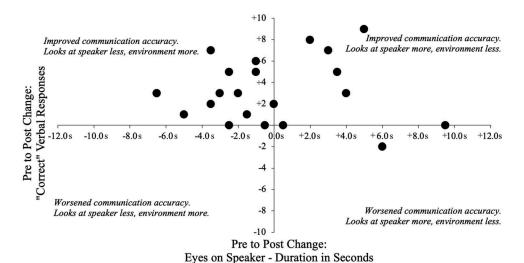


Figure 3. Scatterplot showing pretest to posttest changes in 1) participant communication accuracy and 2) participant eye fixation duration on speakers while they are speaking. Descriptions are provided in italics for each quadrant.

environment longer. One participant decreased their frequency of correct verbal responses in posttest, while looking at the speaker longer. These patterns provide insight into the measurement and assessment of situational awareness in virtual simulations: there is no one "right" way to be situationally aware. Some individuals increased their awareness of events, as measured by communication accuracy, by looking at speakers longer than the environment. Others increased their awareness by looking at the environment to which the speaker is referring, longer than at the speakers themselves. Individual histories likely bear influence on which method is most effective.

The results of this study provide a preliminary basis for assessing communication accuracy and situational awareness during simulated events. Many technical factors likely influenced these results. The variation among these factors was not assessed in the present study; therefore, we consider these limitations which future research should assess: past VR experience among participants, frequency of participant exposure to simulations, duration of simulations, amount of required participant interaction, variation of simulation scenarios used, and number of participants within each simulation. Furthermore, the effects of the full TeamSTEPPS® didactic training on performance can be assessed. Future research can further utilize VR by creating and programming fully interactive virtual environments. An interactive virtual environment would allow participants to pick up and manipulate objects, communicate directly with one another in the same simulation, and interact directly with a simulated virtual patient. Such a virtual environment could expand the analysis from a single individual to that of a group. Findings from such studies would demonstrate even greater external validity, as most operations within HROs occur in groups.

The future of high-precision measurement and assessment of behavior is conceivably tied to VR. VR simulations allow for a high degree of experimental control heretofore unachievable in other simulated scenarios, greater measurement precision of difficult-to-measure responses, greater variety and customization potential of scenarios used, and the emerging ability of simulations to support multiple users at once. As VR technology grows in capability and costs to produce this technology continue to decrease, it has the potential to transition from specialized uses to everyday applications for behavior analysts - from simulating basic lab research for behavior analytic students, to assessment and training applications with individual clients.

Acknowledgments

The authors wish to thank the following individuals for their significant contributions to this study: Christopher Lewis (University of Nevada, Reno Department of Computer Science & Engineering) for his technical assistance in analyzing eye-tracking data and providing on-site support during simulation sessions; Sven Diaz (University of Nevada, Reno Department of Computer Science & Engineering), Hadi Rumjahn and Tyler Goffinet (University of Nevada, Reno @One Digital Media Technology) for their on-site support during simulation sessions; and Alicia Fong (University of Nevada, Reno School of Medicine) for her logistic coordination of simulation sessions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Agency for Healthcare Research and Quality (AHRQ). (2008). Hospital survey on patient safety culture: 2008 comparative database report. Retrieved from http://www.ahrq.gov/qual/ hospsurvey08

Agency for Healthcare Research and Quality (AHRQ). (2012). TeamSTEPPS 2.0. Retrieved from https://www.ahrq.gov/teamstepps/instructor/index.html

Alavosius, M. P., Houmanfar, R. A., Anbro, S. J., Burleigh, K., & Hebein, C. (2017). and crew resource management in high-reliability organizations: A competency framework for measuring behaviors. Journal of Organizational Behavior Management, 37(2), 142–170. doi:10.1080/01608061.2017.1325825

Bijou, S. W., Chao, C. C., & Ghezzi, P. M. (1988). Manual of instructions for identifying and analyzing referential interactions II. The Psychological Record, 38(3), 91-98. doi:10.1007/ BF03395032

Crosswell, L., Porter, L., & Sanders, M. (2018). Out of sight, out of mind? Addressing unconscious brand awareness in healthcare communication. In H. D. O'Hair (Ed.), Risk and health communication in an evolving media environment (pp. 55-77). Waltham, MA: Focal Press.

Endsley, M. R. (1995a). Toward a theory of situational awareness in dynamic systems. Human Factors, 37, 32-64. doi:10.1518/001872095779049543



- Endsley, M. R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors*, 37, 65–84. doi:10.1518/001872095779049543
- Epps, H. R., & Levin, P. E. (2015). The TeamSTEPPS approach to safety and quality. *Journal of Pediatric Orthopedics*, 35(5), S30–S33. doi:10.1097/BPO.00000000000000541
- Hollowell, B., & Lansing, C. R. (2004). Tracking eye movements to study cognition and communication. *The ASHA Leader*, 9(21), 4–5.
- inspireUX. (2010). *Eyetracking metrics for usability studies*. Retrieved from http://www.inspireux.com/2010/06/14/eyetracking-metrics-for-usability-studies/
- Johnson, R. A., Houmanfar, R. A., & Smith, G. S. (2010). The effect of implicit and explicit rules on customer greeting and productivity in a retail organization. *Journal of Organizational Behavior Management*, 30(1), 38–48. doi:10.1080/01608060903529731
- Killingsworth, K., Miller, S. A., & Alavosius, M. P. (2016). A behavioral interpretation of situational awareness: Prospects for organizational behavior management. *Journal of Organizational Behavior Management*, 36(4), 301–321. doi:10.1080/01608061.2016.1236056
- King, H. B., Battles, J., Baker, D. P., Alonso, A., Salas, E., Webster, J., ... Salisbury, M. (2008). TeamSTEPPS: Team strategies and tools to enhance performance and patient safety. In K. Henriksen, J. B. Battles, M. A. Keyes, & M. L. Grady (Eds.), *Advances in patient safety: New directions and alternative approaches (Vol. 3: Performance and tools)*. Rockville, MD: Agency for Healthcare Research and Quality. pp. 5-20.
- Makary, M. A., & Daniel, M. (2016). Medical error The third leading cause of death in the US. *The British Medical Journal*, 353, i2139–i2143. doi:10.1136/bmj.i2139
- Maraccini, A. M., Houmanfar, R. A., Kemmelmeier, M., Piasecki, M., & Slonim, A. D. (2018). An inter-professional approach to train and evaluate communication accuracy and completeness during the delivery of nurse-physician student handoffs. *Journal of Interprofessional Education & Practice*, 12, 65–72. doi:10.1016/j.xjep.2018.06.003
- Sawyer, T., Laubach, V. A., Hudak, J., Yamamura, K., & Pocrnich, A. (2013). Improvements in teamwork during neonatal resuscitation after interprofessional TeamSTEPPS training. *Neonatal Network*, 32(1), 26–33. doi:10.1891/0730-0832.32.1.26
- Sheppard, F., Williams, M., & Klein, V. R. (2013). TeamSTEPPS and patient safety in healthcare. *Journal of Healthcare Risk Management*, 32(3), 5-10. doi:10.1002/jhrm.21099
- Smith, G. S., Houmanfar, R. A., & Denny, M. (2012). Impact of rule accuracy on productivity and rumor in an organizational analog. *Journal of Organizational Behavior Management*, 32(1), 3–25. doi:10.1080/01608061.2012.646839
- Tobii Pro. (n.d.). *Dark and bright pupil tracking*. Retrieved from https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-is-dark-and-bright-pupil-tracking/
- Turner, P. (2012). Implementation of TeamSTEPPS in the emergency department. *Critical Care Nursing Quarterly*, 35(3), 208–212. doi:10.1097/CNQ.0b013e3182542c6c
- Ward, T. A., & Houmanfar, R. A. (2011). Human simulations in behavior analysis (1987–2010): Facilitating research in complex human behavior. *Behavior and Social Issues*, 20, 72–101. doi:10.5210/bsi.v20i0.3707