A Simulation-based Approach to Measuring Team Situational Awareness in Emergency Medicine: A Multicenter, Observational Study

Elizabeth D. Rosenman, MD, Aurora J. Dixon, MA, Jessica M. Webb, MA, Sarah Brolliar, Simon J. Golden MA, Kerin A. Jones, MD, Sachita Shah, MD, James A. Grand, PhD, Steve W. J. Kozlowski, PhD, Georgia T. Chao, PhD, and Rosemarie Fernandez, MD

ABSTRACT

Objectives: Team situational awareness (TSA) is critical for effective teamwork and supports dynamic decision making in unpredictable, time-pressured situations. Simulation provides a platform for developing and assessing TSA, but these efforts are limited by suboptimal measurement approaches. The objective of this study was to develop and evaluate a novel approach to TSA measurement in interprofessional emergency medicine (EM) teams.

Methods: We performed a multicenter, prospective, simulation-based observational study to evaluate an approach to TSA measurement. Interprofessional emergency medical teams, consisting of EM resident physicians, nurses, and medical students, were recruited from the University of Washington (Seattle, WA) and Wayne State University (Detroit, MI). Each team completed a simulated emergency resuscitation scenario. Immediately following the simulation, team members completed a TSA measure, a team perception of shared understanding measure, and a team leader effectiveness measure. Subject matter expert reviews and pilot testing of the TSA measure provided evidence of content and response process validity. Simulations were recorded and independently coded for team performance using a previously validated measure. The relationships between the TSA measure and other variables (team clinical performance, team perception of shared understanding, team leader effectiveness, and team experience) were explored. The TSA agreement metric was indexed by averaging

From the Division of Emergency Medicine, University of Washington (EDR, SB, SS, RF), Seattle, WA; the Department of Psychology (AJD, JMW, SJG, SWJK) and the Department of Management (GTC), Michigan State University, East Lansing, MI; the Department of Emergency Medicine, Wayne State University (KAJ), Detroit, MI; and the Department of Psychology, University of Maryland (JAG), College Park, MD.
Received April 24, 2017; revision received June 29, 2017; accepted July 12, 2017.
This study was funded by grants from the Agency for Healthcare Research and Quality (1R18HS020295 [RF] and 1R18HS022458-01A1 [RF]). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation of, approval of, or decision to submit the manuscript.
EDR's institution has received grant money from the Agency for Healthcare Research and Quality and the Department of Defense to conduct research conceived and written by RF. EDR reports personal payment from Physio-Control, Inc. for speaker fees. JMW's institution has received grant money from the Department of Defense to conduct research conceived and written by RF. SB is supported by grants from the Agency for Healthcare Research and Quality and the Department of Defense to conduct research conceived and written by RF. SJG's institution has received grant money from the Agency for Healthcare Research and Quality to conduct research conceived and written by RF. SS's institution has received grant money from the Department of Defense to conduct research conceived and written by RF. SS has received funding personally from Sonosite Fujifilm, Ultralinq (Bothell, WA), for consulting work evaluating ultrasound technology and from Group Health/Kaiser Permanente (Seattle, WA) for teaching ultrasound as a consultant. JAG's institution has received grant money from the Department of Defense to conduct research conceived and written by RF and JAG. SWJK's institution has received grant money from the Agency for Healthcare Research and Quality and the Department of Defense to conduct research conceived and written by RF. SWJK's institution has received grant money from NASA to conduct research conceived and written by SWJK. GTC's institution has received grant money from the Agency for Healthcare Research and Quality and the Department of Defense to conduct research conceived and written by RF. GTC's institution has received grant money from NASA to conduct research conceived and written by SWJK. RF's institution has received grant money from the Agency for Healthcare Research and Quality, the Department of Defense, and Physio-Control, Inc., to conduct research conceived and written by RF. RF reports personal payment from Physio-Control, Inc., for speaker fees.
The other authors have no potential conflicts to disclose.
Supervising Editor: Robert Cloutier, MD.
Address for correspondence and reprints: Rosemarie Fernandez, MD; e-mail: fernanre@uw.edu
ACADEMIC EMERGENCY MEDICINE 2018;25:196–204.

 ISSN 1553-2712
doi: 10.1111/acem.13257

the pairwise agreement for each dyad on a team and then averaging across dyads to yield agreement at the team level. For the team perception of shared understanding and team leadership effectiveness measures, individual team member scores were aggregated within a team to create a single team score. We computed descriptive statistics for all outcomes. We calculated Pearson's product-moment correlations to determine bivariate correlations between outcome variables with two-tailed significance testing ($p < 0.05$).

Results: A total of 123 participants were recruited and formed three-person teams (n = 41 teams). All teams completed the assessment scenario and postsimulation measures. TSA agreement ranged from 0.19 to 0.9 and had a mean (±SD) of 0.61 (±0.17). TSA correlated with team clinical performance ($p < 0.05$) but did not correlate with team perception of shared understanding, team leader effectiveness, or team experience.

Conclusions: Team situational awareness supports adaptive teams and is critical for high reliability organizations such as healthcare systems. Simulation can provide a platform for research aimed at understanding and measuring TSA. This study provides a feasible method for simulation-based assessment of TSA in interdisciplinary teams that addresses prior measure limitations and is appropriate for use in highly dynamic, uncertain situations commonly encountered in emergency department systems. Future research is needed to understand the development of and interactions between individual-, team-, and system (distributed)-level cognitive processes.

Emergency medical systems depend on teams to effectively function and adapt within dynamic, complex situations. Highly adaptive teams can meet variable patient- and system-level demands by modifying their cognitive, affective, and behavioral processes.[1] Failure to act proactively and react appropriately can threaten patient safety and result in diagnostic errors.[2] This ability to adapt depends on the development of team cognitive structures necessary to support a coordinated response to, or in anticipation of, change.

Situational awareness supports the cognitive and behavioral processes necessary for resilient, adaptive healthcare systems.[3,4] Situational awareness is defined as the "perception of the elements in the environment within the volume of time and space, the comprehension of their meaning, and the projection of the status in the near future."[5] Within this definition are three domains:

1. Level 1—perception of the environment, which involves information gathering from multiple sources including patient-related data, electronic health records, and other clinicians.
2. Level 2—comprehension of the meaning of information and events, which requires synthesis, interpretation, and prioritization of available data to create an understanding of the current state of the system.
3. Level 3—projections of future states, which involves predicting one or more possible trajectories to allow for contingency planning for high probability events and anticipation of next steps

Situational awareness is conceptualized at the individual, team, and organizational levels. Team situational awareness (TSA) is the shared understanding among team members that facilitates team coordination and task performance.[6] TSA supports dynamic decision making and adaptability in unpredictable, time-pressured situations, and has been shown to improve team effectiveness in emergency response teams.[7]

Measuring TSA in healthcare teams is important for several reasons. TSA measures support effective team training by providing quantifiable indicators of TSA that inform performance feedback.[8] As Salas et al. note, "without quantifiable indicators of TSA it is hard to articulate what constitutes good situation awareness." TSA measures are also important for evaluating training effectiveness, identifying areas of weakness (e.g., clinical situations where TSA development is particularly challenging), determining team competency, and identifying when failures to develop TSA contribute to poor team performance. Finally, TSA measures may help improve system design by allowing researchers to test system-level interventions (e.g., process changes, equipment design, physical layout) that support development of TSA.

Measuring TSA in healthcare teams is challenging.[4,8] Measures are often indirect, using subjective self-report questionnaires[9–11] or behavioral assessments[12] that measure the actions supporting the development of situational awareness. While such measures are useful, they do not capture the cognitive processes underlying clinical decision making. The most frequently reported direct measure of situational awareness, the situation awareness global assessment technique (SAGAT), involves executing a simulation and "freezing" activity at set points to ask learners questions pertaining to the three levels of situational awareness.[13] This type of measure, known as an "in-task probe," requires interrupting the simulation and diverting the attention of the learners,[6] which may negatively influence learning, especially when applied at the team level.[14] Several studies in emergency medicine (EM) and critical care used the SAGAT;[15–17] however, most measured the situational awareness of individuals, focusing on accuracy relative to a "criterion standard" rather than similarity with other team members. In studies of team cognition, both similarity and accuracy have been shown to impact team performance.[18] Given the diagnostic uncertainty and complexity encountered in EM, the shared understanding among team members, rather than accuracy, may be a more appropriate target for assessment.

The objective of this study was to develop and evaluate a novel approach to simulation-based TSA measurement in interprofessional EM teams. Existing conceptual frameworks and empirical research in team cognition and team science support this approach. This work will provide EM educators, researchers, and system operations experts with a broader understanding of the importance of TSA and its contribution to team performance. Furthermore, it offers an approach to assessing TSA that takes into consideration the nature of EM teams.

METHODS

Study Design

We performed a multicenter, prospective, simulation-based observational study to evaluate validity evidence supporting an approach to TSA measurement in interprofessional EM teams. Institutional review boards both at the University of Washington and at Wayne State University approved the study.

Participants and Setting

We recruited EM resident physicians (n = 41), EM nurses (n = 41), and medical student participants (n = 41) from two different urban, academic institutions, the University of Washington (Seattle, WA) and Wayne State University (Detroit, MI). EM residents were all in postgraduate year 2, 3, or 4 of training and in good standing within their respective residency programs. Emergency department (ED) nurses were all certified in Advanced Cardiac Life Support (ACLS) and had a minimum of 1 year of nursing experience in an urban ED. Medical students were recruited from years 1 through 4 and were in good academic standing. Participation was voluntary for all participants, and all participants received U.S. $25 compensation for their participation. Simulations were conducted from June 2014 through October 2014 at the Kado Family Clinical Skills Center (Wayne State University) and the Washington, Wyoming, Alaska, Montana, and Idaho (WWAMI) Institute for Simulation in Healthcare (University of Washington).

Study Protocol

Eligible participants provided written consent and were assigned to three-member teams (one resident physician, one nurse, one medical student) on the basis of schedule availability. The teams completed a simulated clinical event of a patient in respiratory distress, progressing to cardiac arrest. Prior to starting the simulation, all team members completed demographic information and experience questionnaires using Qualtrics software. Immediately following the simulation, and prior to any discussion or debriefing, all team members completed several measures including: 1) a TSA measure, 2) a team perception of shared understanding measure, and 3) a team leader effectiveness measure. All simulations were recorded using two stationary cameras. Team members wore individual microphones to enhance audio capture. Recordings were later reviewed and coded in duplicate by trained raters using a team clinical performance measure.

Medical Resuscitation Scenario

All simulations were performed using a Laerdal SimMan human patient simulator and a nurse (actor) with a partially scripted role designed to provide information (e.g., test results) at specific times. We modified an event-based scenario of a patient with shock physiology complicated by cardiac arrest.[19] Briefly, the patient presented from a skilled nursing facility to the ED with dyspnea and altered mental status and rapidly progressed to respiratory distress and ventricular fibrillation. The patient had return of spontaneous circulation in response to appropriate treatment as outlined by 2015 American Heart Association ACLS guidelines. If the team did not perform critical actions within a certain time frame, the nurse actor prompted appropriate treatment so all simulations ended in return of spontaneous circulation.

The need to modify our original simulation required collecting new validity evidence.[20] The scenario, including clinical content, timing of events, behavioral triggers, and nurse script, was reviewed by EM physicians (n = 3) and EM nurses (n = 2). The scenario was then piloted with code teams (n = 3) and team members were interviewed. Team members were specifically queried about scenario representativeness, the psychological fidelity of behavioral prompts and clinical events, and the nursing script content. Together these steps provided evidence of content and response process validity for the scenario. Finally, the reliability of the scenario was assessed throughout the study period to ensure that all clinical cues and triggers were presented in a standardized manner.

Outcome Measures

TSA Measure. We developed a TSA measure (Data Supplement S1, available as supporting information in the online version of this paper, which is available at https://doi.org/onlinelibrary.wiley.com/doi/10.1111/acem.13257/full) designed to be completed by all team members immediately following the resuscitation scenario. The TSA measure focused on clinical information that was not role specific (i.e., all team members could be expected to know the answer). The questions emphasized the projection domain of situational awareness, (Level 3) as projecting from the current state to predict future events is thought to be most critical for performance in highly dynamic settings.[21,22] Questions were reviewed by EM attending physicians (n = 3) and piloted within an interprofessional team to ensure readability and representativeness of items. Pilot team members were interviewed and the TSA measure was refined based on feedback. The final measure included seven multiple-choice questions. The TSA agreement metric was indexed by averaging the pairwise agreement for each dyad on a team (1 = agreement, 0 = no agreement), yielding the proportion of agreement for each dyad (physician–nurse, physician–medical student, nurse–medical student) and then averaging across dyads to yield the proportion of agreement for the team as a collective. Dyads where a single data point was missing were considered to be in disagreement for that item.

We also queried the team members about their perception of the team's shared understanding using a single multiple-choice question (Data Supplement S2, which is available at https://doi.org/onlinelibrary.wiley.com/doi/10.1111/acem.13257/full). We analyzed this item separately from the TSA measure as it reflected self-perception only.

Team Leader Effectiveness Measure. Effective team leaders distribute and interpret information to build coherence and a shared understanding of the situation.[23] We therefore evaluated whether our TSA measure correlated with team member perception of team leader effectiveness using a six-item questionnaire (Data Supplement S2). Questions were designed and reviewed by team science and clinical subject matter experts to ensure questions reflected team processes important for medical resuscitation.[24] Results from factor analyses of these items showed the six leader effectiveness items to load on one factor, so we averaged these items to form a leadership effectiveness score.

Team Clinical Performance Measure. To assess team clinical performance we adapted an evidence-based clinical team performance measure previously shown to have good inter-rater reliability and validity.[19] Example outcome measure formats included time to completion of a task (e.g., medication administration), presence or absence of a behavior (e.g., performing an airway assessment), and the number of behaviors occurring simultaneously. Minor modification of the original measure was required to ensure the individual items reflected the modified scenario. Team performance measure items contained multiple formats (e.g., whether or not a behavior occurred, time to completion of an activity); therefore, we first standardized the data (i.e., converted to z-scores). A team's overall clinical performance score was the weighted average of all items. In our previous work we describe the process for determining item weights.[19] Trained raters consisting of two EM physicians independently coded video recordings of simulations for team performance using Noldus Observer XT software. We calculated Cohen's kappa for each recording and averaged the values to give an overall measure of reliability (Cohen's κ = 0.86).[25]

Data Analysis

Statistical analyses were calculated using SPSS Version 24 (IBM Corp.). Collective team experience was indexed using a multi-item composite that combined different types of experiences across team members. Discrete items assessing the resident months spent in the ED, resident months on critical care rotations, nurse experience in emergency nursing (years), nurse experience in critical care nursing (years), medical student year, simulation experience (number of sessions), and the number of resuscitations led were subjected to a principal components analysis. The analysis yielded a seven-component solution that was then subjected to a varimax rotation. Corresponding component weights from this solution were applied to each of the discrete experience items (i.e., multiplied) and summed to create the experience composite. Thus, the composite appropriately accounted for the unique contribution of each experience item.

The normality of variable distributions was examined for skewness and kurtosis. We computed descriptive statistics (mean and standard deviation [SD] or median and interquartile range as appropriate) for all outcomes. We calculated Pearson's product-moment correlations to determine bivariate correlations between outcome variables with two-tailed significance testing ($p < 0.05$).

RESULTS

A total of 123 participants were recruited (resident physicians [n = 41], nurses [n = 41], and medical students [n = 41]). Participants were 38% (n = 47) male with a mean (±SD) age of 32 (±8.0) years. Residents were in postgraduate years 2 (61%; 25/41), 3 (32%; 13/41), and 4 (7%; 3/41). Nurses reported a mean (±SD) of 13 (±8.8) years nursing experience, with a mean (±SD) of 10.6 (±7.0) years in EM nursing. Of the 41 teams, 10% (4/41) had a medical student in the first or second preclinical year. All 41 teams completed the assessment scenario, TSA measure, team perception of shared understanding item, and team leader effectiveness measure. Because the team clinical performance measure was modified for this study, we reestablished a high level of interrater reliability for team performance (Cohen's $\kappa = 0.86$).[25]

Descriptive analyses and outcome correlations are shown in Table 1. TSA agreement ranged from 0.19 to 0.9 and had a mean (±SD) of 0.61 (±0.17). TSA positively correlated with team clinical performance ($p < 0.05$) as predicted by the team cognition literature.[8] Team experience, team leader effectiveness, and team perception of shared understanding did not correlate with TSA. Team leader effectiveness positively correlated with team perception of shared understanding ($p < 0.001$).

DISCUSSION

Team situational awareness supports adaptive teams and is critical for high reliability organizations such as healthcare systems.[26] The skills underlying TSA can be trained, and simulation provides an effective mechanism to support the complex environments necessary to develop TSA. Simulation can also provide a platform for research aimed at understanding team cognition, e.g., TSA, and its impact on the healthcare system. These simulation-based training and research efforts require a conceptually sound and valid approach to TSA measurement. This study provides a novel method for assessing TSA in interprofessional teams that addresses prior measure limitations and is appropriate for use in highly dynamic, uncertain situations commonly encountered in ED systems.

Our methods were based on a previously described process for measure development in simulation-based assessments[27] and a measurement validation approach outlined by Cook et al.[28] Our approach to TSA measurement addresses several challenges associated with applying SAGAT and other TSA assessment methodologies to EM teams. First, and most importantly, our methodology involves postsimulation assessment rather than in-simulation assessments. While proponents of the SAGAT suggest "in-task" probes do not impact performance,[13,29] research in highly dynamic team

Table 1
Descriptive Statistics and Correlations of Study Measures

Measures	Mean (±SD)	Median (IQR)	1	2	3	4
1. Team clinical performance	0.47 (±0.09)	–				
2. Team experience composite*	–	–0.48 (3.59)	–0.13			
3. Team situational awareness agreement	0.61 (±0.17)	–	0.32†	0.27		
4. Perception of shared understanding‡	3.78 (±0.46)	–	–0.05	0.26	0.25	
5. Team leader effectiveness‡	3.48 (±0.46)	–	–0.03	0.07	0.11	0.55§

N = 41 teams.
*Experience composite scores did not fit a normal distribution.
†$p < 0.05$.
‡Outcomes are individual team member perceptions averaged across a team.
§$p < 0.001$
IQR = interquartile range.

settings suggest otherwise.[30] EM team tasks are highly interdependent, requiring high levels of adaptive coordination for effective team performance. Studies suggest that interruptions under such conditions decrease team members' attention and increase stress levels.[30] Furthermore, the effect of interruptions can accumulate over time and thus be dependent on the length of the simulation and number of interruptions.[31]

Second, our approach focused on similarity, rather than accuracy, and thus the correlation we found between team clinical performance and TSA reflected the value of a shared understanding among team members, rather than correctness. Research in other areas of team cognition recognizes that more than one correct cognitive model is possible in a given situation, and thus comparing results to a single criterion standard may be inadequate.[32] This is pertinent to the diagnostic uncertainty within EM, and for this reason we purposely designed a simulated scenario that contained some ambiguity. The mean agreement across teams was 61%. As we found no other studies evaluating TSA similarity in resuscitation teams, it is difficult to interpret this result within a larger body of literature, but we expect this agreement across team members is lower than it would be in a more straightforward clinical scenario with a "correct" answer.

We provide several types of validity evidence[28] supporting a TSA measure delivered after a simulated clinical event. We had EM subject matter experts review the simulation and all measures to establish content validity. We piloted the TSA measure on an interprofessional team and interviewed team members to determine alignment of items with TSA and clinical content, thus providing response process validity evidence. This allowed us to refine our measure and move forward with data collection to evaluate for evidence of construct validity (i.e., relationship with other variable).

Our results demonstrated a positive correlation between team clinical performance and TSA agreement. We feel that this relationship represents our strongest evidence of validity. The link between TSA and performance is clearly supported by both theoretical and empirical work.[8,33,34] Additionally, our team clinical performance measure was previously validated for use with the simulated scenario used in this study.[19] We do note that the correlation between team clinical performance and TSA was weak to moderate.[35] As with the TSA agreement discussed, the strength of the correlation between our TSA measure and team clinical performance likely reflects the ambiguous nature of the simulation and the purposeful lack of diagnostic certainty. Studies evaluating correlations of performance and *accuracy* of situational awareness in *individuals* demonstrate similar strengths of association.[9,16] We extended this work by evaluating correlations of performance and *similarity* of situational awareness in interprofessional *teams*. The interprofessional nature of our teams may also influence the degree of shared awareness. Future work could focus on delineating the components of TSA that should be shared versus complimentary to maximize team performance, care delivery, and system effectiveness.

We did not find a correlation between our TSA measure and team perception of shared understanding. This is not necessarily surprising, as self-assessments are largely thought to be inaccurate.[36] Social science and cognition researchers suggest that perceptions of situational awareness likely reflect an individual's confidence level in his/her understanding, rather than providing a true measure of situational awareness.[13,37] Of note, we also did not find a correlation between team experience and TSA. This is also not surprising, since it is well known that expert individual performance does not necessarily translate into expert team performance. In other studies demonstrating correlations between experience and situational awareness, the situational awareness measure focused on accuracy and had less within team experience variability.[38]

While we present a novel approach to measuring TSA, it is important to recognize that situational awareness is a complex concept and appreciating this complexity requires familiarity with the larger concept of team cognition. The cognitive capability of a team is not simply the sum of the cognition of the individuals, but rather it depends on how that information is structured and accessed. Related constructs include team mental models and transactive memory systems. Team mental models refer to shared knowledge and transactive memory systems refer to the collective awareness of how specialized knowledge is distributed within the team.[39]

Of the team cognition constructs, situational awareness has gained the most traction in healthcare and is included in the Agency for Healthcare Research and Quality's TeamSTEPPS curriculum.[40] The increasing focus on aspects of team cognition is a move in the right direction; however, these complex cognitive processes are often oversimplified.[41] People discuss "losing situational awareness" as if it is a tangible entity.[42] In reality, situational awareness is not well defined. We do not know whether it is a proxy for underlying processes, the cognitive state supporting these processes, or the resulting outcome. Maintaining distinction between individual-, team-, and system-level constructs is also important and often a source of error in team research.[43,44] While we attempted to evaluate team-level factors most likely to impact TSA, such as team experience, it is possible there are individual-level factors not evaluated in this study that could influence TSA outcomes. More work is needed to understand exactly how to represent, measure, and improve TSA in healthcare.

This study has several strengths. First, it is a multicenter study. While certain clinical guidelines, such as those provided by the American Heart Association, are adopted nationally, there are regional and institutional practices that reflect system-specific issues. It is therefore important to evaluate measures and assessment approaches across multiple institutions in different geographic regions. Second, we were able to utilize an existing simulation-based assessment platform (simulation + performance measure) supported by evidence of validity.[19] Third, our approach is conceptually consistent with the nature of EM teams, the dynamic nature of their tasks, and the uncertainty inherent in their environment.[24] Finally, our subject size exceeds what is typically found in healthcare team–level research, where we found studies assessing a team-level component of situational awareness to have sample sizes ranging from three to 10.[30,38,45,46]

LIMITATIONS

There are several limitations to this study. First, our assessment methodology involves a simulated task. While we and others recommend simulation for team-based cognitive assessments,[8,47] simulation-based research may not capture all of the subtle environmental and system cues that inform cognitive processes. However, simulation affords the opportunity to include all members of the team in the assessment. In real ED clinical events, providers often have to move quickly to the next patient care task and do not have time to thoughtfully complete measures. Rigorously designed simulations also provide a standardized platform for assessment that is not possible in the actual clinical environment. Second, our measures could be subject to hindsight bias, as team members have access to more information at the end of the simulation than at specific time points of interest during the simulation. We were careful to design questions that reflected time points in the simulation after which further information was not disclosed. We also note that this potential limitation would likely have less of an impact on TSA measure similarity than accuracy. Third, we acknowledge that we did not assess role-specific knowledge or TSA accuracy as we felt TSA similarity is an important starting point for cognitive study in EM teams. We also note that subjects' TSA assessment responses may be influenced by multiple individual- and team-level factors that were not measured within the framework of this study. Our study focused on agreement as an attempt to measure TSA independent of individual knowledge. Assessing TSA across multiple different clinical events could help to further mitigate the impact of clinical knowledge and individual experience on outcomes. Finally, we note that this research provides preliminary evidence for our TSA measurement approach. Continued evaluation of this method should be pursued across multiple different simulations, team types, and settings to further establish its validity, generalizability, and boundary limitations.

CONCLUSION

Team cognition is critical for effective EM team and system performance. This study provides a novel approach to assessing an important team-level construct that is supported by multiple types of validity evidence. Rigorously designed simulation-based research, whether conducted in laboratories or in situ, provides an opportunity to train the team behaviors that support cognitive processes and assess the content and structure of team situational awareness. Future research is needed to understand the development of and interactions between individual-, team-, and system (distributed)-level cognitive processes.

We thank Ms. Heather Brooks-Buza (Wayne State University) and Ms. Beth Sobba (University of Washington) for their assistance implementing study simulations. We also thank Ann Line, DVM, for her assistance with project coordination. Finally, we express our heartfelt thanks to the medical students, nurses, and residents at both institutions for their participation in this study.

References

1. Baard SK, Rench TA, Kozlowski SW. Performance adaptation: a theoretical integration and review. J Manage 2014;40:48–99.
2. National Academy of Sciences, Engineering, and Medicine. Improving Diagnosis in Health Care. Washington, DC: The National Academies Press, 2015.
3. Wright MC, Taekman JM, Endsley MR. Objective measures of situation awareness in a simulated medical environment. BMJ Qual Saf 2004;13:i65–71.
4. Lowe DJ, Ireland AJ, Ross A, Ker J. Exploring situational awareness in emergency medicine: developing a shared mental model to enhance training and assessment. Postgrad Med J 2016;92:653–8.
5. Endsley MR. Toward a theory of situation awareness in dynamic systems. Hum Factors 1995;37:32–64.
6. Wildman JL, Salas E, Scott CP. Measuring cognition in teams: a cross-domain review. Hum Factors 2014;56:911–41.
7. van der Haar S, Segers M, Jehn K, Van den Bossche P. Investigating the relation between team learning and the team situation model. Small Gr Res 2015;46:50–82.
8. Salas E, Prince C, Baker DP, Shrestha L. Situation awareness in team performance: implications for measurement and training. Hum Factors 1995;37:123–36.
9. Plant JL, van Schaik SM, Sliwka DC, Boscardin CK, O'Sullivan PS. Validation of a self-efficacy instrument and its relationship to performance of crisis resource management skills. Adv Health Sci Educ 2011;16:579–90.
10. Mullan PC, Macias CG, Hsu D, Alam S, Patel B. A novel briefing checklist at Shift Handoff in an emergency department improves situational awareness and safety event identification. Pediatr Emerg Care 2015;31:231–38.
11. Walia J, Qayumi Z, Khawar N, et al. Physician transition of care: benefits of I-PASS and an electronic handoff system in a community pediatric residency program. Acad Pediatr 2016;16:519–23.
12. Russ S, Hull L, Rout S, et al. Observational teamwork assessment for surgery feasibility of clinical and nonclinical assessor calibration with short-term training. Ann Surg 2012;255:804–9.
13. Endsley MR. Measurement of situation awareness in dynamic systems. Hum Factors 1995;37:65–84.
14. Salmon PM, Stanton NA, Walker GH, et al. Measuring situation awareness in complex systems: comparison of measures study. Int J Ind Ergon 2009;39:490–500.
15. Siegel NA, Kobayashi L, Dunbar-Viveiros JA, et al. In situ medical simulation investigation of emergency department procedural sedation with randomized trial of experimental bedside clinical process guidance intervention. Simul Healthc 2015;10:146–53.
16. Cemalovic N, Scoccimarro A, Arslan A, et al. Human factors in the emergency department: is physician perception of time to intubation and desaturation rate accurate? Emerg Med Australas 2016;28:295–9.
17. Stratton A, Furey A, Hogan M. The effect of sleep deprivation on a resident's situational awareness in a trauma scenario. J Orthop Trauma 2014;28:e160–3.
18. Lim BC, Klein KJ. Team mental models and team performance: a field study of the effects of team mental model similarity and accuracy. J Organ Behav 2006;27:403–18.
19. Fernandez R, Pearce M, Grand JA, et al. Evaluation of a computer-based educational intervention to improve medical teamwork and performance during simulated patient resuscitations. Crit Care Med 2013;41:2551–62.
20. Validity Messick S. in: Linn RL, ed. Educational Measurement. vol. 3. New York: American Council on Education and Macmillan, 1989:13–103.
21. O'Brien K, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. Ergonomics 2007;50:1064–91.
22. Sulistyawati K, Wickens CD, Chui YP. Prediction in situation awareness: confidence bias and underlying cognitive abilities. Int J Aviat Psychol 2011;21:153–74.
23. Kozlowski SW, Gully SM, Salas E, Cannon-Bowers JA. Team leadership and development: theory, principles, and guidelines for training leaders and teams. in: Beyerlein MM, Johnson DA, Beyerlein ST, eds. Advances in Interdisciplinary Studies of Work Teams: Team Leadership. vol. 3. New York: Elsevier Science/JAI Press, 1996:253–91.
24. Fernandez R, Kozlowski SW, Shapiro MJ, Salas E. Toward a definition of teamwork in emergency medicine. Acad Emerg Med 2008;15:1104–12.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
26. Wilson KA, Burke CS, Priest HA, Salas E. Promoting health care safety through training high reliability teams. BMJ Qual Saf 2005;14:303–9.
27. Grand JA, Pearce M, Rench TA, et al. Going DEEP: guidelines for building simulation-based team assessments. BMJ Qual Saf 2013;22:436–48.
28. Cook DA, Zendejas B, Hamstra SJ, Hatala R, Brydges R. What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. Adv Health Sci Educ 2014;19:233–50.
29. Endsley M. A methodology for the objective measurement of pilot situation awareness. AGARD, Situational Awareness in Aerospace Operations 9 (SEE N 90-28972 23-53). 1990.
30. Morgan P, Tregunno D, Brydges R, et al. Using a situational awareness global assessment technique for interprofessional obstetrical team training with high fidelity simulation. J Interprof Care 2015;29:13–9.
31. Weigl M, Antoniadis S, Chiapponi C, Bruns C, Sevdalis N. The impact of intra-operative interruptions on surgeons' perceived workload: an observational study in

elective general and orthopedic surgery. Surg Endosc 2015;29:145–53.
32. Mathieu JE, Heffner TS, Goodwin GF, Cannon-Bowers JA, Salas E. Scaling the quality of teammates' mental models: equifinality and normative comparisons. J Organ Behav 2005;26:37–56.
33. Endsley MR. Situation Awareness Global Assessment Technique (SAGAT). Paper presented at Aerospace and Electronics Conference, 1988
34. Rosen MA, Bedwell WL, Wildman JL, et al. Managing adaptive performance in teams: guiding principles and behavioral markers for measurement. Hum Resour Manage R 2011;21:107–22.
35. Colton T. Statistics in Medicine. Boston: Little, Brown and Company, 1974.
36. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence - A systematic review. JAMA 2006;296:1094–102.
37. Fracker ML. Measures of Situation Awareness: Review and Future Directions. DTIC Document; 1991.
38. Hogan MP, Pace DE, Hapgood J, Boone DC. Use of human patient simulation and the Situation Awareness Global Assessment Technique in practical trauma skills assessment. J Trauma 2006;61:1047–52.
39. DeChurch LA, Mesmer-Magnus JR. The cognitive underpinnings of effective teamwork: a meta-analysis. J Appl Psychol 2010;95:32–53.
40. TeamSTEPPS® Curriculum Materials (content last reviewed April 2017). Agency for Healthcare Research and Quality. Available at: http://www.ahrq.gov/teamstepps/curriculum-materials.html. Accessed Apr 24, 2017.
41. Dekker S, Hollnagel E. Human factors and folk models. Cogn Technol Work 2004;6:79–86.
42. Dekker SW. The danger of losing situation awareness. Cogn Technol Work 2015;17:159–61.
43. Dionne SD, Chun JU, Hao CY, et al. Article quality and publication impact via levels of analysis incorporation: an illustration with transformational/charismatic leadership. Leadersh Q 2012;23:1012–42.
44. Klein KJ, Kozlowski SW. From micro to meso: critical steps in conceptualizing and conducting multilevel research. Organ Res Methods 2000;3:211–36.
45. Crozier MS, Ting HY, Boone DC, et al. Use of human patient simulation and validation of the Team Situation Awareness Global Assessment Technique (TSAGAT): a multidisciplinary team assessment tool in trauma education. J Surg Educ 2015;72:156–63.
46. Gardner AK, Kosemund M, Martinez J. Examining the feasibility and predictive validity of the SAGAT tool to assess situation awareness among medical trainees. Simul Healthc 2017;12:17–21.
47. Gaba DM, Howard SK. Situation awareness in anesthesiology. Hum Factors 1995;37:20–31.

Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13257/full

Data Supplement S1. Team situational awareness measure.

Data Supplement S2. Team perception of shared understanding measure and team leader effectiveness measure.