Measuring situation awareness in emergency settings: a systematic review of tools and outcomes

Simon Cooper1,2
Joanne Porter3
Linda Peach4
1School of Nursing and Midwifery, Monash University, Berwick, Victoria, Australia; 2School of Nursing and Midwifery, University of Brighton, Brighton, UK; 3School of Nursing and Midwifery Monash University, Gippsland, VIC; 4School of Nursing and Midwifery, Monash University, Melbourne, VIC, Australia

Background: Nontechnical skills have an impact on health care outcomes and improve patient safety. Situation awareness is core with the view that an understanding of the environment will influence decision-making and performance. This paper reviews and describes indirect and direct measures of situation awareness applicable for emergency settings.

Methods: Electronic databases and search engines were searched from 1980 to 2010, including CINAHL, Ovid Medline, Pro-Quest, Cochrane, and the search engine, Google Scholar. Access strategies included keyword, author, and journal searches. Publications identified were assessed for relevance, and analyzed and synthesized using Oxford evidence levels and the Critical Appraisal Skills Programme guidelines in order to assess their quality and rigor.

Results: One hundred and thirteen papers were initially identified, and reduced to 55 following title and abstract review. The final selection included 14 papers drawn from the fields of emergency medicine, intensive care, anesthetics, and surgery. Ten of these discussed four general nontechnical skill measures (including situation awareness) and four incorporated the Situation Awareness Global Assessment Technique.

Conclusion: A range of direct and indirect techniques for measuring situation awareness is available. In the medical literature, indirect approaches are the most common, with situation awareness measured as part of a nontechnical skills assessment. In simulation-based studies, situation awareness in emergencies tends to be suboptimal, indicating the need for improved training techniques to enhance awareness and improve decision-making.

Keywords: nontechnical skills, situation awareness, teamwork, emergency, acute care

Introduction

This paper follows our review of nontechnical skill measures in emergency care1 in which we identified core nontechnical skills, including an approach to measurement of situation awareness (SA). In this paper, we review measures of SA that have been tested in acute care/emergency settings and describe the outcomes.

Core aspects of nontechnical skills include teamwork, leadership, decision-making, and SA, with measures available for leadership, teamwork, personality, behavior, and SA. Nontechnical skills do have an impact on health care outcomes2 and do improve patient safety.3 For a number of years, SA has been a core focus for aircraft cockpit crews,4 the military,5 and for driver instruction,6 with the view that an understanding of the environment will influence personal and team performance.

Wickens defines SA as: “the continuous extraction of information about a dynamic system or environment, the integration of this information with previously acquired knowledge to form a coherent mental picture, and the use of that picture in directing
further perception of, anticipation of, and attention to future events."

The theoretical underpinnings are that SA is the operator's internal model of the environment from which they can decide what actions to take, making SA the main precursor to decision-making. However, SA is influenced by long-term memory, working memory, and attention focus. Experience therefore leads to the development of internal models of the environment, reducing the demand on working memory and enabling rapid or "recognition primed" decisions and therefore automaticity. Individuals seek information relevant to achieve a goal, which can lead to reformation of a goal, whilst the ability to switch between goal-directed and data-directed behavior has an influence on SA. However, whilst high levels of SA may increase the probability of a "good" decision, this is not guaranteed, because poor decisions are apparent despite high levels of SA. Further, in higher workload situations, SA and attention focus are likely to be reduced, especially in critical "time-bound" situations.

Situation awareness measurement approaches vary, and include direct experimental techniques, which are the most common approaches. The Situation Awareness Global Assessment Technique (SAGAT) is one such approach where, during suspension of the task, probes and queries are used to ascertain levels of SA. Verbal protocols are another approach where participants are asked to "think aloud" during a task, and there are self-administered tests, such as the Situational Awareness Rating Technique, that produce an overall rating of participants' perceived level of SA. Concerns are raised here, given that such measurements may simply reflect confidence in SA and not actual SA. Observer ratings may also indicate only the behavior of participants and not the internal processing of information.

Situations are dynamic, and the aim must be to identify when a situation has changed and demands an altered response. In the emergency setting, this is complicated by time constraints in a rapidly changing setting, with the need to measure and judge based on the "obtainable ideal." The applications of SA measurement are diverse, and include training development, error analysis, behavioral prediction, teamwork, and insights into issues such as automation and workload (with suggestions that increases in automation free up cognitive space with improved SA). Measures of SA have also been found to be sensitive to differences that are not identified in other measures of performance whilst retaining a correlation with measures of knowledge.

In the following section, we review both individual and team measures of SA, with a particular focus on the SAGAT in emergency settings. In order to capture all possible measures, "emergency situations" are broadly defined as acute medical events occurring in any clinical or simulated setting.

Methods
A search of the literature was conducted to locate and review instruments that rated SA measures applicable for acute emergency care settings in English from 1980 to 2010. Four databases were used, ie, CINAHL, Ovid Medline, Pro-Quest, Cochrane, and the search engine, Google Scholar. The primary search terms were "teamwork", "situation awareness", and "simulation". CINAHL and Ovid Medline were searched independently using the primary search terms followed by a search of related MeSH terms and linking papers. The selected papers were then reviewed by the two lead authors, leaving a total of 55 studies that were analyzed and synthesized using the Oxford evidence levels and the Critical Appraisal Skills Programme guidelines for assessing the quality and rigor of original research. In consultation with the authorship team, papers were included if they reported measures of SA that were applicable to emergency situations and undertaken in any country. Papers not available in English were excluded. Following a primary search, papers that did not meet the inclusion criteria by title or abstract were excluded, and the remaining full papers were reviewed by two authors.

One hundred and thirteen papers were initially identified and then reduced to 55 following title and abstract review. The final selection included 14 papers drawn from the fields of emergency medicine, intensive care, anesthetics, and surgery. Ten of these discussed four general nontechnical skill measures, which included SA, and four reviewed and used SAGAT (Figure 1).
Results

Team and individual measures of nontechnical skills that include SA are listed as follows, including the Team Emergency Assessment Measure (TEAM), Anesthetic Non-Technical Skills (ANTS), Non-Technical Skills for Surgeons (NOTSS), and Non-Technical Skills (NOTECHS) (Table 1). The authors tend to use Endsley’s definition of SA:

“Situation awareness refers to an individual’s “perception of the elements in the environment within the volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”.”

Team Emergency Assessment Measure

TEAM was developed from earlier versions of the Emergency Team Dynamics tool and has been extensively tested for reliability and validity in the settings of resuscitation and patient deterioration. Intended as an emergency team nontechnical skills measure, the eleven items are grouped under three categories of leadership, teamwork, and task management, and rated on a scale of 0–4. A twelfth item is included as a global “overall” rating of team performance. Content validity is high, with a content validity index of 0.96 and an internal consistency (Cronbach’s alpha) of 0.89. For interrater reliability, no individual item assessments were measured, but the mean intraclass correlation coefficient for the 11 items was 0.60. In relation to SA measures, under the teamwork category, there are two elements relating to SA (perception), “The team monitored and reassessed the situation”, and SA (projection), “The team anticipated potential actions”. TEAM and guidance on how to use the tool can be found at http://medicareemergencyteam.com/.

Anesthetic Non-Technical Skills

Developed by a team of industrial psychologists, ANTS is an observed behavior rating scale using a rating of 1 (poor) to 4 (good). Trained observers rate four key nontechnical skill categories, including “task management”, “team working”, “decision-making”, and “situation awareness”. The latter has three elements, i.e., “gathering information”, “recognizing and understanding”, and “anticipating”. Based on rating of eight simulated anesthetic scenarios by 50 trained observers, internal consistency between the elements using Cronbach’s alpha was satisfactory, ranging from 0.79 to 0.86. However, interrater reliability was lowest in the SA category (mean within-group interrater agreement indices, $r_wg$ 0.56), indicating the difficulty in judging SA from an external perspective. Further work by Yee et al indicates similar levels of interrater reliability.

Analysis of critical incident studies suggests that the ANTS framework is also relevant to work in intensive care. Reader et al found that task management accounted for most of the nontechnical skills factors contributing to critical incidents in the intensive care unit, closely followed by SA factors.

Non-Technical Skills for Surgeons

Yule et al describe the initial development of this tool, identifying five nontechnical skill categories, i.e., “situation awareness”, “decision-making”, “leadership”, “communication and teamwork”, and “task management”. The latter category has been removed in some versions of the tool. SA elements include “gathering information”, “understanding information”, and “projecting and anticipating future state”. In testing the reliability of the tool with 44 surgeons rating six experimental sessions, NOTSS was found to have a consistent internal structure and good interrater reliability for the categories of “communication and teamwork” ($r_wg$ 0.70) and “leadership” ($r_wg$ 0.72), but poor reliability for “situation awareness” ($r_wg$ 0.51). Intraclass correlation coefficients indicated high agreement, ranging from 0.95 to 0.99.

Non-Technical Skills

Originally developed by Flin et al for aircraft pilots, this tool has been adapted for the measurement of nontechnical skills in operating department assistants, scrub nurses, anesthetists, and surgeons. Vigilance/SA is measured and, for operating department assistants and nurses, includes measures of monitoring, awareness, and anticipation, and for surgeons and anesthetists, includes measures of active communication, monitoring, and awareness. In these adaptations, communication appears to be based on a top-down one-way interpretation, with only surgeons and anesthetists being rated on communication in a crisis and excluded from measures of anticipation. In contrast, one of the key ANTS measures of SA is “anticipation”, suggesting that the adaptations of NOTECHS are incomplete. Further, Sevdalis et al reviewed the internal consistency of a revised version of NOTECHS in observations of surgical teams in simulated settings, and identified a Cronbach’s alpha of 0.77–0.87 and that SA/vigilance had the least consistency/reliability overall for subdivisions of professional groups (0.66–0.85).

The above measures use observational ratings of SA that, as discussed, may only indicate the behaviors of participants and not the internal processing of information. Further, interrater reliability and the internal consistency of SA measures tend to be lower than for other nontechnical skill categories.
suggesting that SA is a challenging concept to rate, with significant interpretative elements. The solution to this dilemma is direct measurement techniques, such as SAGAT.

**Situation Awareness Global Assessment Technique**

Using SAGAT, three levels of SA can be measured, covering respondents’ “perception” and “understanding” of the situation and their “prediction” of future events.

Because SAGAT is not based on self or observer ratings of SA, it is considered a more objective measure with greater reliability and validity.

With reference to the situation to be measured (eg, a resuscitation attempt), SA questions are developed by a team of experts to assess the elements of perception, comprehension, and projection. Using goal-directed task analysis, experts identify the goals and subgoals associated with a work task, and the decisions required to meet these goals, in order to produce questions related to the three levels of SA (see Figure 2 for an example of this process in a cardiac deterioration scenario). When assessing SA in simulation exercises, the session can be stopped at a random point (ie, a “freeze”) and participants asked the standardized set of questions in a “quick fire” rapid sequence to ascertain their immediate “gut reaction” response. Freezes during the scenario are the standard approach, but Wright et al suggest that this could impact on successful performance. In qualitative analysis relating to our first exploratory patient deterioration management study, we found indications that the freeze had negatively affected participants’ clinical performance and altered their SA focus as the scenario developed. Whilst this may be beneficial in developing SA skills, the interruption in the flow of the scenario did not match reality and would not be applicable in a “real” clinical event. Subsequently, we have always asked SA questions immediately following each scenario, a process that is achievable in the clinical setting. Further, in relation to measurement of SA levels, we have subdivided perception into physiologic perception (eg, “what is the blood pressure?”) and global perception (eg, “what was on the wall near the patient?”) measures (see Figure 2). The latter enables measurement of SA in the wider scene, with the expectation that global awareness will be low in emergency situations, especially when managed by novices.

In a range of exploratory studies using SAGAT and other measures, we aimed to identify how health professionals manage acute episodes of patient deterioration in short 8-minute, primary responder, simulation exercises. Low SA scores were identified, averaging around 53% across scenarios and groups. For example, the SA of student nurses was 59% (range 38%–82%) and was lowest in the comprehension domain and global SA domain. Student midwives averaged 54% (range 40%–70%), and had lowest scores in the physiologic perception domain, mainly because they failed to measure vital signs. Registered nurses working in rural settings had similar levels of SA at 50%, with low levels of global and physiologic perception (again possibly related to failure to measure vital signs), but with an expected higher level of comprehension and anticipation (projection) than student nurses. A second study of rural hospital nurses identified the same overall rating of SA (50%) and the same ranking of the separate domains, lowest to highest being global perception, physiologic perception, comprehension, and projection. Further, in this study, higher SA scores were significantly correlated with a younger age group. These nursing studies identify that SA tends to be low in simulated emergency settings, possibly due to high levels of anxiety. For example, clinical skill performance...
does decrease as the patient deteriorates and the situation becomes more demanding.36

In summary, SAGAT has been tested in a variety of fields, including driving, piloting, and air traffic control,15 while Hogan et al16 have used this approach in trauma simulations. Endsley39 reports that the SAGAT has produced results that are consistently valid and reliable, and that the technique is sensitive to system manipulations, automation manipulations, differences in expertise, and operational concepts across a variety of domains. However, there are concerns that SAGAT is more representative of memory than of the SA level, a view that is countered by Endsley, who argues that working memory is an essential component of SA and therefore the two are intertwined.

Discussion

This systematic review was limited to peer-reviewed papers published in English in the last 20 years. Additional tools may be available in the gray and unpublished literature. We identified a small number of tools and techniques that we identified as being applicable for the measurement of SA in emergency settings; however, these tools were tested in heterogeneous populations, so their degree of generalizability may be questioned.

It is apparent that nontechnical skills do have an impact on health care outcomes2 and do improve patient safety.1 SA is core with the view that an understanding of the environment will influence personal and team performance.7 Evidence from surgery29 and anesthesia23 indicates that an increase in SA will improve medical staff performance, whilst Mishra et al40 identified that, in surgeons performing laparoscopic cholecystectomy, technical skills (error rate) was negatively correlated with their level of SA. SA is one of the main precursors to decision-making,4 but can degrade with fatigue and stress and be affected by interruptions and distractions.

SA measurements vary with direct measurement techniques such as SAGAT and the observational ratings included in larger nontechnical skills assessments such as the TEAM. Observational ratings may only indicate behavior and not necessarily SA itself; however, there is clearly a trade-off between the feasibility of a measure and its validity and reliability. SAGAT is a valid technique and has been described in the health care literature and in emergency settings. It is feasible, but users should consider when to ask SA questions, bearing in mind the study or training objective. Random freezes of scenarios may impact on performance,15 but will be a more accurate record of SA. In the clinical setting, this will not be possible and SA should be addressed at the end of an event.

There is a need to develop effective training strategies to improve SA. Core approaches recommended41 include:

- video-recorded high stakes, high fidelity simulation
- an accurate and feasible measure of performance that includes task and team management and SA domains

Table 1 Situation awareness measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Rigor/outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Emergency Assessment Measure (TEAM)</td>
<td>An observational rating scale designed initially as a resuscitation team NTS measure. TEAM includes two elements related to SA (projection and perception).</td>
<td>A content validity index of 0.96 and an internal consistency of 0.89. Mean intraclass correlation coefficient for the 11 items was 0.60. Internal consistency ranged from 0.79 to 0.86. Interrater reliability was lowest in the SA category (0.56). Good interrater reliability for the categories of “communication and teamwork” (0.70) and “leadership” (0.72) but lower reliability for SA (0.51). Intraclass correlation coefficients were 0.95–0.99. Internal consistency was 0.77–0.87. SA had the least consistency overall for subdivisions of professional groups (0.66–0.85).</td>
</tr>
<tr>
<td>Anesthetic Non-Technical Skills (ANTS)</td>
<td>An observational rating scale designed to measure anesthetists’ NTS in four categories, one of which is SA.</td>
<td></td>
</tr>
<tr>
<td>Non-Technical Skills for Surgeons (NOTSS)</td>
<td>An observational rating scale designed to measure surgeons’ NTS in five categories, one of which was SA.</td>
<td></td>
</tr>
<tr>
<td>Non-Technical skills (NOTECH5)</td>
<td>An observational rating scale designed to measure NTS for ODPs, scrub nurses, anesthetists, and surgeons. Vigilance/situation awareness is measured in slightly different ways for each professional group.</td>
<td></td>
</tr>
<tr>
<td>Situation Awareness Global Assessment Technique (SAGAT)</td>
<td>A direct measurement technique for SA measuring “perception”, “understanding” and “prediction” of future events. Goal-directed task analysis used to develop the goals and decisions associated with the work task to produce questions related to the levels of SA.</td>
<td>In nursing and midwifery trials in simulated settings, low SA scores have been identified (53%). Higher SA scores were correlated with a younger age group.</td>
</tr>
</tbody>
</table>

Abbreviations: SA, situation awareness; NTS, nontechnical skills; ODPs, operating department assistants.
video debriefing and feedback techniques that incorporate participant self-reflection and review of performance outcomes

post event clinical review, eg, resuscitation team debriefings.

Advanced techniques for recording events are now available, that aid debriefing and inform SA, and ultimately patient safety. For example, in 2013, we are running a trial with paramedics and nurses using eye tracking video glasses that record the global scene and gaze pathway in emergencies. Henneman et al used the same technique to assess whether clinicians checked patient identity bands before administering medication.

Conclusion

A range of direct and indirect techniques for measuring SA are available. In the medical literature, indirect approaches are the most common, with SA measured as part of a non-technical skills assessment. Such approaches may lack the rigor and accuracy of measuring SA directly with techniques such as SAGAT. In simulation-based studies, SA in emergencies tends to be suboptimal, indicating the need for improved training techniques to enhance awareness and improve decision-making, with ultimate benefits to patient safety.

Disclosure

The authors report no conflicts of interest in this work.

References