

Objective measures of situation awareness in a simulated medical environment

M C Wright, J M Taekman, M R Endsley

Qual Saf Health Care 2004;13(Suppl 1):i65-i71. doi: 10.1136/qshc.2004.009951

One major limitation in the use of human patient simulators is a lack of objective, validated measures of human performance. Objective measures are necessary if simulators are to be used to evaluate the skills and training of medical practitioners and teams or to evaluate the impact of new processes or equipment design on overall system performance. Situation awareness (SA) refers to a person's perception and understanding of their dynamic environment. This awareness and comprehension is critical in making correct decisions that ultimately lead to correct actions in medical care settings. An objective measure of SA may be more sensitive and diagnostic than traditional performance measures. This paper reviews a theory of SA and discusses the methods required for developing an objective measure of SA within the context of a simulated medical environment. Analysis and interpretation of SA data for both individual and team performance in health care are also presented.

applied in a medical simulation environment. Advantages and disadvantages of various measurement methods are described. We present a theory of situation awareness (SA) and, as an example, discuss the implications of SA within anaesthesia. We propose that direct measures of SA through the situation awareness global assessment technique (SAGAT) may provide an effective objective measure of individual and team performance in a patient simulation environment. We describe the methods required for the use and analysis of SAGAT for patient simulation applications.

MEASURING HUMAN MACHINE SYSTEMS

In evaluating performance, we are typically concerned with understanding the ability of the individual medical care provider to perform various tasks while utilising the tools and medical care systems at his or her disposal. This combined human system performance seeks to evaluate not the individual per se, but the degree to which the tool or devices either aid or undermine that performance. This type of evaluation is extremely important in evaluating the design of new technology for the medical care setting. In addition, we may be interested in evaluating individual or team performance in order to assess the quality of new training regimens.

A number of different types of measures have been used to assess human machine systems in a variety of work environments. These include direct and indirect measures of performance, mental workload measures, and a range of analytic measures of specific aspects of performance such as movement or communication.¹ Examples include:

- Direct performance measures—measures of outcome or human machine "score", time on task, error rate, degree of error (such as deviation from planned path);
- Indirect performance measures—subjective ratings of performance including both self ratings and outside observer ratings;
- Mental workload measures—subjective ratings of workload, secondary task measures (better performance on a secondary task implies greater spare capacity suggesting lower workload), physiological measures;

Abbreviations: GDTA, goal directed task analysis; SA, situation awareness; SAGAT, situation awareness global assessment technique; SART, situation awareness rating technique

Human patient simulators can be a valuable resource for studying issues related to medical error. Using simulators, researchers are able to study the effects of various adverse events on the performance of surgeons, anaesthesiologists, other medical practitioners, and nurses. This provides a great improvement to the "hit or miss" method of training and post hoc study of adverse events associated with the treatment of real patients, where adverse events are likely to occur only rarely and may never be encountered during the training period. In addition, human patient simulators can be used to show how new equipment design and new processes or procedures may affect the performance of clinicians, and thus, the potential resulting outcomes for patients.

Using simulators to evaluate the effects of training, medical equipment design, or changes in process or procedures ultimately requires valid measures that can provide data to support conclusions regarding the effectiveness of these interventions. Unfortunately, few measures have been developed for quantifying the performance of medical professionals using simulators. In addition, the validation process for these measures can be arduous and is currently an important topic of discussion within the medical simulation community.

This paper reviews existing methods of measuring human machine systems that may be

See end of article for authors' affiliations

Correspondence to: Melanie C Wright, PhD, Department of Anesthesia, Box 3094, Duke University Medical Center, Durham, NC 27710 USA; melanie.wright@duke.edu

Key messages

- Objective measures of human system performance are needed for use in evaluations conducted using human patient simulators.
 - Situation awareness (SA) is defined as a person's perception of elements in the environment, comprehension of that information, and the ability to project future events based on this understanding.
 - SA is a critical component in decision making for medical practitioners.
 - Objective measures of SA can be more sensitive and more diagnostic than traditional measures of human performance.
 - Objective measures of SA such as situation awareness global assessment technique (SAGAT) may provide an effective means of assessing both individuals and teams in a human simulation environment.
 - The use of SAGAT requires detailed analysis of the task to be studied to identify SA requirements in order to develop appropriate SA queries.
 - Results of evaluations using objective measures of SA with human patient simulators may be used to improve medical practitioner training and medical equipment design.
- Task analytic measures—analysis of eye tracking data, time and motion studies, time spent on various components of a task, communications analysis.

Direct performance measures are difficult to define for assessing the performance of clinicians in a simulated patient environment as there may be many possible solutions to a particular problem. Possible measures might include success or failure in a given task or scenario or time to identify a specific problem or adverse event that has been preprogrammed into a simulation. Researchers have used written tests,² observed errors,³⁻⁵ problem detection and diagnosis time,⁶ task completion time,³ and measurement of simulator variables such as effect site concentration of drugs.⁷ Unfortunately, these measures provide little evidence as to why poor performance may have occurred. In addition, overall performance measures often reflect only the outcome of the task or event. Therefore, they may not identify errors or misconceptions that are resolved before the task is complete or that are not reflected in the outcome of the event (for example, certain drug errors).⁸

Subjective^{7,9} and secondary task¹⁰ measures of workload have also been used in patient simulations. While workload measures are useful in identifying situations where the clinician may suffer from overload, this is only one aspect of a task that is likely to influence performance. In some cases measures of workload and performance dissociate. For example, poor performance is sometimes associated with passive monitoring tasks where the individual is not actively involved in performing a task, often termed the out of the loop problem.¹¹ While workload may be low, performance can suffer because the individual is not aware of changes that may be occurring (for example, the patient has taken a turn for the worse).

Indirect performance measures have been used successfully in human simulation environments. Several researchers have used observer ratings to assess the performance of clinicians in a simulated environment.¹²⁻¹⁴ Gaba *et al* found that this method resulted in slightly better interrater

reliability when used to evaluate technical performance (for example, motor performance associated with tasks such as chest compression) compared to behavioural performance (such as decision making and team communication skills).¹³ Gaba describes several problems associated with this type of measure including the high cost of using multiple experienced raters and high interindividual variability that may occur between raters.¹⁵

SITUATION AWARENESS

Cognitive constructs such as attention and mental workload are useful in the study of human performance when formally defined and integrated into testable theories. Situation awareness is another such construct. Situation awareness can be thought of as an internal mental model of the current state of an individual's environment. Endsley has formally defined SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future".¹⁶ This definition breaks the concept of SA into three distinct levels including: (1) Level 1—perception of the environment, (2) Level 2—comprehension of the meaning of this information, and (3) Level 3—projection of events or actions in the future based on this perception and comprehension.

In Endsley's model of SA (figure 1), SA is shown as a construct that is distinct from decision making and the performance of actions.^{17,18} However, SA precedes decision making and is one important component of dynamic decision making. Other factors including both task or system factors and individual factors as shown in the model also influence this process. For example, two practitioners may have the same SA, but choose different courses of action based on their prior clinical experience. Personality characteristics such as risk aversion or system constraints may also affect their decisions. SA as defined in this model includes only that portion of a person's knowledge that pertains to the state of a dynamic environment. Background knowledge, experiences, and established rules are static knowledge sources that fall outside of the definition of SA, although they may influence the development of SA. For example, preconceptions based on previous experiences can direct an individual's attention, affecting the formation of SA. Another important detail of this model is that SA is continuously changing as the environment changes, either due to the decisions and actions of the individual or due to other outside influences.

Klein specifies four reasons why the phenomenon of SA is important in the study of human work,¹⁹ these include:

1. SA appears to be linked to performance. This assertion has obvious face validity since it is expected that the more relevant information a worker has about a situation the more adaptive their responses will be.
2. Limitations in SA may result in errors. If needed information is not available or is not correctly interpreted, whether due to failures of memory or attention or due to system failures, then clearly there is an increased likelihood of errors.
3. SA may be related to expertise. Klein raises examples of classic research where experienced physics researchers have been shown to classify physics problems differently than novices.¹⁹
4. SA is the basis for decision making in most cases. This characteristic of SA is an integral part of both Klein's recognition primed decision model¹⁹ and Endsley's model of SA.¹⁷

Research in aviation and other environments where SA has been measured supports Klein's four reasons for studying SA.

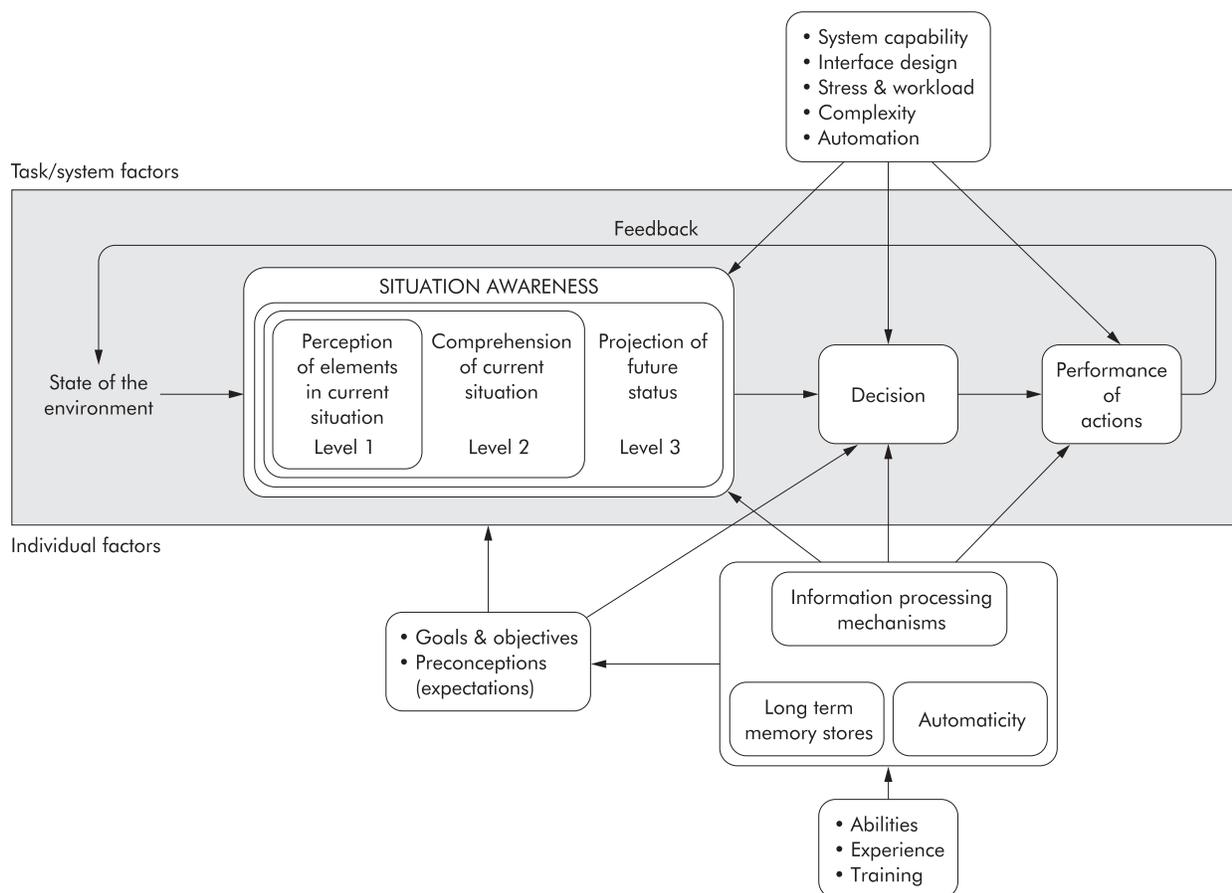


Figure 1 Model of situation awareness.¹⁷ Reprinted with permission from *Human factors*, vol 37, No 1, 1995. Copyright 1995 by the Human Factors and Ergonomics Society. All rights reserved.

Measures of SA have correlated with performance in aviation.^{20, 21} Bell and Lyon found that fighter pilots with lower observer ratings of SA during a combat scenario had a greater number of decision errors than pilots with more highly rated SA.²² In addition, SA measures have been shown to be sensitive to both task difficulty and experience in aviation and in power plant operations.^{23, 24}

More interesting, however, is the research that has shown measures of SA to be sensitive to differences that were not reflected in performance measures. For example, Endsley conducted a study comparing the SA of pilots using a new avionics system with that of pilots using the old system.²⁵ Pilots subjectively believed the new system to be better, but mission performance measures showed no differences. Endsley used SAGAT, a direct SA measurement technique in which pilots are asked questions regarding their perception, comprehension, and projection of the current situation during a simulation freeze to evaluate the new system. She found that the new system provided pilots with better SA regarding knowledge of enemy aircraft location and other critical factors compared to the old system.

In another example, concern about the effects of a new form of air traffic control known as “free flight” on the ability of the air traffic controllers to track and monitor aircraft led to a comparison of performance and SA between the new and old systems.²⁶ Performance tests with the new system showed trends toward performance differences regarding separation errors, although the results were not significant. SAGAT measures in this experiment were able to provide more diagnostic detail showing that controllers were aware of significantly fewer aircraft under free flight conditions (level

1 SA), that controllers had a significantly reduced understanding of what was happening in the traffic situation (level 2 SA), and that controllers had reduced knowledge of where aircraft were going (level 3 SA). These studies suggest that measures of SA can have diagnostic powers beyond measures of performance that may be predictive of performance problems or errors that are not seen within the limited sensitivity, scope, or time involved in a laboratory study.

SITUATION AWARENESS IN HUMAN PATIENT SIMULATION ENVIRONMENTS

Human patient simulators show great promise for the improvement of both training methods and equipment design in a wide range of medical applications. The practice of anaesthesia is a good example.²⁷ The environment of the anaesthetist is highly dynamic, complex, uncertain, risky, and subject to intense time pressure.²⁸ The study of cognitive processes of humans has shown that this type of environment is very challenging to the human as a decision maker.²⁷ Humans are poor at performing tasks under low stress and low workload conditions (for example, monitoring tasks) and are also poor at performing tasks under high stress and high workload conditions. These types of working conditions affect the individual’s ability to focus attention in the appropriate areas. In addition, task complexity, dynamism, and uncertainty are likely to interact to reduce a person’s ability to make sound decisions.

For an anaesthesiologist, developing and maintaining SA can be a difficult task because of the large amount of information present in a complex and dynamic environment.

The following examples apply Endsley's three levels of SA²⁹ to the tasks of an anaesthesiologist:

- Level 1 SA—Perception of the elements in the environment. The first step in achieving SA is to perceive the status, attributes, and dynamics of relevant elements in the environment. For an anaesthesiologist, these may include awareness of patient vital signs such as heart rate, blood pressure, oxygen saturation, breathing rate, laboratory values, and hospital studies and may also include awareness of drugs and the patient's level of consciousness. Awareness of the actions of other team members such as the surgeon and nurses and awareness of equipment function (including potential problems) also are critical elements of level 1 SA for the anaesthesiologist.
- Level 2 SA—Comprehension of the current situation. Comprehension of the situation is based on a synthesis of the separate level 1 elements. Level 2 SA involves understanding the significance of objects and events in the environment and combining this data to form a holistic picture of the environment in light of one's goals. For example, anaesthesiologists will synthesise and integrate information regarding patient physical signs and patient information to identify the most probable cause in a complex differential diagnosis. They will understand the significance of a sudden drop in heart rate, based on knowledge of a surgeon's recent procedures and other vital signs, to know if it represents an expected and temporary event or a serious problem.
- Level 3 SA—Projection of future status. The highest level of situation awareness is to project the future actions of elements in the environment. This is achieved through knowledge of the status and dynamics of elements in the environment and comprehension of the situation (level 1 and level 2 SA). Anaesthesiologists with a high degree of level 3 SA will be able to project the response of the patient to anticipated drug administration and physician actions, including changes in vital signs such as breathing rate and oxygen saturation. This type of projection is very important in allowing them to be proactive and not just reactive.

In the medical environment, the level of SA possessed by a practitioner may be critical to the outcome of the patient. The measurement of SA allows researchers to determine whether clinicians have good SA in relation to the task environment. This could help in the identification of performance problems and error mechanisms (possibly induced by equipment with a poor user interface, poor arrangements of equipment in the surgical setting, or poor communication and teamwork among the surgical staff). Measurement of SA may be used as a method of evaluating training by identifying areas of deficiency (that is, areas where individuals fail to attain the needed levels of SA). The results can be used to improve training and education of medical practitioners. A measure of SA may also lead to improvements in system design (such as improved equipment design or process changes) that will support medical practitioners in attaining a high level of SA which will allow them to make the best decisions, resulting in better patient outcomes.

MEASURING SITUATION AWARENESS IN A SIMULATED PATIENT ENVIRONMENT

Measures of SA can be either indirect, such as subjective ratings, or direct, usually through probing workers with direct questions regarding the task. Subjective measures can be either self rating techniques such as the situation awareness rating technique (SART) by Taylor³⁰ or they may be ratings made by outside observers. Endsley has suggested that self ratings most likely reflect a measure of subjects'

confidence level regarding their SA, rather than providing a true measure of their SA.⁸ In the case of outsider observer ratings, these can only be based on behaviours and verbalisations of the subjects; therefore, they can not determine whether specific information has been stored and processed internally.

Gaba *et al* suggest the use of real time probes to measure the SA of anaesthesiologists using patient simulations.³¹ They suggest that queries be raised by actors within the scenario (such as doctors or nurses) and the participant's responses can be used to assess their level of SA. Endsley, however, states that such queries may bias the results of the experiment by artificially directing the subject's attention to certain parameters.⁸

Endsley presents SAGAT as a method of directly measuring SA.⁸ In this method, a simulation is frozen at various points in time and workers are asked questions designed to assess their level 1, level 2, and level 3 SA. The answers are then compared to the real situation according to the simulated computer database and experts' interpretations of the meaning of that data to provide an objective measure of SA. One important aspect of SAGAT is the development of queries for the experiment. This requires a fairly in-depth study of an individual's role to identify the SA requirements and appropriate phrasing of questions. Methods such as Goal Directed Task Analysis (GDTA) may be used to identify task goals, related decisions, and finally the SA requirements that are needed to make the decisions that allow operators to meet their goals.³²

SAGAT has been criticised regarding the perceived intrusiveness of freezes in a simulation to collect SAGAT data, and the degree to which it reflects memory and as such is limited.³³ Research collected to date does not support these concerns. Several studies have shown that a temporary freeze in a simulation to collect SAGAT data does not impact performance.^{8 34} Other studies have used SAGAT in some trials, but not others in order to examine whether SAGAT interferes with participant performance^{32 35-37} and found no differences indicating either better or worse performance when SAGAT is being used. Studies have also found that subjectively participants who used SAGAT appeared to adjust to the technique quite well and are able to return to the action fairly readily after a short freeze to collect SAGAT data.^{32 38}

In terms of an over-reliance on memory, Endsley³² provides a detailed examination of this issue finding that memory is central to SA and that SAGAT taps into the working memory stores where information is integrated and processed to form the ongoing dynamic situation representation, as well as long term memory stores that feed SA. SAGAT in fact overcomes the many problems of retrospective report of past mental events, due to its use of freezes to collect these mental impressions of the situations as immediately as possible and without other intervening events to disrupt memory. Endsley^{8 34} found that pilots' ability to report their SA via SAGAT was unaffected by how long after the freeze the question was asked, testing intervals from around 20 s to up to 6 min, showing a lack of memory decay for this information due to integration of both working and long term memory stores. The SAGAT methodology is specifically designed to tap into the SA resident in human memory as effectively as possible, using cued recall queries administered as quickly and immediately as possible.

The application of SAGAT to medical applications is limited. Zhang and colleagues used a variation on SAGAT to compare two anaesthesia displays in a patient simulation. They found significant differences in SA due to type of display for level 1 and level 2 SA for some of the scenarios

tested.⁶ One problem with their method, however, was that they used a very small subset of queries (four level 1 queries, two level 2 queries, and two level 3 queries), such that during the course of the experiment participants may be able to predict the questions that would be asked for any given simulation freeze.

The validity of SAGAT has been established in other environments including piloting, driving, and air traffic control.^{8, 32} In general, these results should generalise to the human simulation environment. However, with the exception of the work by Zhang *et al*, we are aware of no published applications of direct measures of SA in patient simulation scenarios.

THE SITUATION AWARENESS GLOBAL ASSESSMENT TECHNIQUE

In order to measure SA, appropriate queries must be developed for assessing level 1, level 2, and level 3 SA. A systematic approach is taken to identify first the goals of medical professionals in the work environment and then the appropriate information requirements to meet those goals. One such approach is a GDTA. For a GDTA, experts are interviewed to identify first the high level goals associated with the task, and then, the sub-goals. A tree structure is created of goals and sub-goals. Further interviews are conducted to identify key decisions for each sub-goal and the SA requirements (at all three levels) that are needed to make those decisions (figure 2). These requirements are then used to develop SA queries for SAGAT. Figure 3 provides an example of a section of a GDTA and some related SA queries an anaesthesia task based on a preliminary GDTA of certified registered nurse anaesthetists. By basing the analysis on goals rather than specific tasks, information requirements are identified that are independent of current technology; thus, the information can be generalised more easily for future efforts.

For the measurement of SA using SAGAT in a simulated patient environment, the patient scenario is frozen at random points during the trial and participants are required to answer SAGAT queries in such a way that they can not see any important perceptual details of the simulation. Endsley³² recommends that the number and timing of simulation freezes and queries meet the following four criteria:

1. The timing of SAGAT stops will be randomly determined.
2. A SAGAT stop will not occur within the first 3–5 min of an experimental trial.

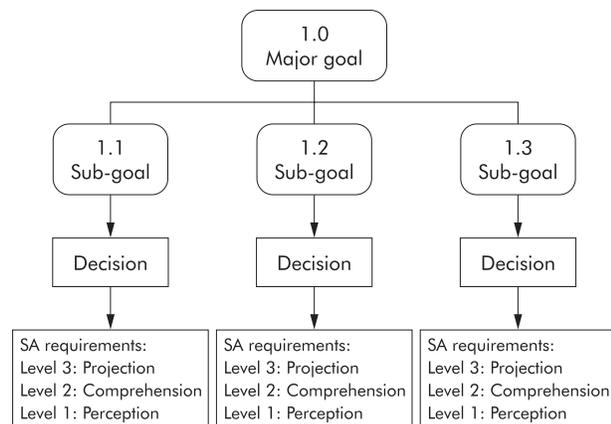


Figure 2 Goal directed task analysis structure including decisions and SA requirements.

3. SAGAT stops will not occur within 1 min of each other.
4. Over the course of the experiment, at least 30 samplings will be collected per SA query (across subjects and trials) for each experimental condition.

The SAGAT responses are scored as correct or incorrect within tolerance bands determined by experts in the area being evaluated, rather than based on absolute error. The frequency of correct responses is tabulated across each query within each experimental condition before further data analyses are completed. The SAGAT data can be analysed both as a composite measure (total SA score based on the number of correct responses to all SA queries) and on an individual query basis. An analysis of individual queries helps provide diagnostic information regarding what types of information or what levels of SA may have been more or less affected by any experimental manipulations such as training conditions or equipment setup.

PRACTICAL APPLICATIONS OF SITUATION AWARENESS MEASURES

Measures of SA such as SAGAT have been used in other domains to both diagnose problems and identify potential solutions including the design of supporting equipment or displays and the development of training programmes. For example, based on the SAGAT results of the air traffic controller evaluation mentioned previously,²⁶ researchers designed a display which provided enhanced information on flight paths for aircraft in transition states as a way of compensating for the lower SA observed.³⁹ With the new display controllers were found to be over three times more likely to be correct in understanding whether aircraft were conforming to their advisories, showing improved SA and ability to perform the task.³⁹

Endsley *et al*⁴⁰ used SAGAT to identify critical differences between experienced and inexperienced general aviation pilots, and found numerous problems with the novice pilots' ability to take in key information, deal with distractions and high workload, monitor effectively, and to understand perceived information and project future events. Based on this research, a set of computer based training modules was designed to build some basic skills underlying SA for new general aviation pilots.⁴¹ These modules included training in time sharing or distributed attention, checklist completion, ATC communications, intensive preflight planning and contingency planning, and SA feedback training which were all found to be problems for new pilots. In tests with low time general aviation pilots, the training modules were generally successful in imparting the desired skills. Some improvements in SA were found in follow on simulated flight trials, but the simulator was insensitive to detect flight performance differences.

Similar research and training development has been conducted for army platoon leaders.^{42, 43} The resulting computer based training programme sought to help build up the mental models and schema that are needed for pattern recognition to produce situation understanding and projection. The training programme also taught skills related to building SA through team communications and contingency planning. In initial testing with cadets performing exercises at the Royal Norwegian Naval Academy, trained cadets were more likely to correctly refuse to attack a refugee camp than untrained cadets, indicating better SA.⁴³ In addition, trained cadets indicated that they spent more mental effort developing level 3 SA and determining how to best meet their goals.

TEAM SITUATION AWARENESS

The theory of SA can be extended to include team environments, such as would be encountered in a surgical



Figure 3 Sample section of a preliminary goal directed task analysis and related SAGAT queries for the role of a certified registered nurse anaesthetist.

setting. Endsley describes team SA as "...the degree to which each team member possesses the SA required for his or her responsibilities".¹⁷ Team members have individual SA requirements and in some cases these requirements overlap, resulting in shared SA requirements. Cooke refers to shared knowledge between team members in two ways: (1) complementary shared knowledge in which the team members have knowledge that does not overlap but is complementary, resulting in the needed team knowledge; and (2) common shared knowledge in which team members share the same knowledge.⁴⁴⁻⁴⁵ A team can be considered to have high team SA when all of the individuals on the team possess the SA (whether complementary or shared) required for their respective roles.

An objective measure of SA such as SAGAT can provide unique insight into team performance within simulated medical environments as well as individual performance. Queries can be designed to assess specific SA requirements for each team member role. More importantly, however, responses to queries related to common SA requirements can be compared across team members, identifying SA differences between team member roles. In addition, specific responses can be compared to determine whether the same responses (correct or incorrect) are made across team member roles. This type of analysis can provide diagnostic

information regarding the source of breakdowns in team SA. For example, common incorrect responses may be indicative of problems that affect the entire team in a similar way (such as poorly designed information display). Alternatively, a mix of correct and incorrect responses or different incorrect responses across team member roles may be indicative of breakdowns in team coordination.

CONCLUSIONS

The potential benefit of patient simulators for training and evaluating medical practitioners and medical equipment is widely recognised.¹⁵⁻⁴⁶ However, methods for measuring the performance of practitioners and the effectiveness of equipment are lacking. The direct measurement of level 1, level 2, and level 3 SA can provide important information regarding the perceptual and cognitive processes of medical practitioners. Future research is needed to determine the SA requirements for a variety of medical tasks or roles (such as anaesthesiology) and to validate measures of SA such as SAGAT within the patient simulation environment. A valid measure of SA may ultimately provide a training tool, in which feedback is given to trainees regarding their performance so that they develop better skills in attaining SA. Such a measure may help designers make important design decisions such as choosing between competing

display designs for monitoring patient vital signs. Measurement of team SA may ultimately lead to training and equipment design that supports better coordination between medical teams which will also lead to higher quality health care.

Authors' affiliations

M C Wright, J M Taekman, M R Endsley*, Department of Anesthesiology and the Duke University Human Simulation and Patient Safety Center, Duke University Medical Center, Durham, North Carolina, USA

*SA Technologies, Inc, Marietta, Georgia, USA

Work attributed to, and funded by the Department of Anesthesiology, Duke University Medical Center, Durham, NC 27710 and SA Technologies, Inc., Marietta, GA.

Competing interests: MRE is the President of SA Technologies. SA Technologies markets software and user guides for the SAGAT.

REFERENCES

- Wilson JR, Corlett EN. *Evaluation of human work: a practical ergonomics methodology*. New York: Taylor and Francis, 1990.
- Morgan PJ, Cleave-Hogg D, McLroy J, et al. Simulation technology: a comparison of experiential and visual learning for undergraduate medical students. *Anesthesiology* 2002;**96**:10–16.
- Merry AF, Webster CS, Weller J, et al. Evaluation in an anaesthetic simulator of a prototype of a new drug administration system designed to reduce error. *Anaesthesia* 2002;**57**:256–63.
- Schwid HA, Rooke A, Carline J, et al. Evaluation of anesthesia residents using mannequin-based simulation: a multi-institutional study. *Anesthesiology* 2002;**97**:1434–44.
- Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance. *Ann Surg* 2002;**236**:458–64.
- Zhang Y, Drews FA, Westenskow DR, et al. Effects of integrated graphical displays on situation awareness in anaesthesiology. *Cognition, Technology & Work* 2002;**4**:82–90.
- Syroid ND, Agutter J, Drews FA, et al. Development and evaluation of a graphical anesthesia drug display. *Anesthesiology* 2002;**96**:565–74.
- Endsley MR. Measurement of situation awareness in dynamic systems. *Hum Factors* 1995;**37**:65–84.
- Herndon CN, Weinger MB, Smith BE, et al. Use of task analysis to evaluate the effects of fatigue on performance during simulated anesthesia cases. *Anesthesiology* 1998;**89**:1180A.
- Byrne AJ, Sellen AJ, Jones JG. Errors on anaesthetic record charts as a measure of anaesthetic performance during simulated critical incidents. *Br J Anaesth* 1998;**80**:58–62.
- Endsley MR, Kiris EO. The out-of-the-loop performance problem and level of control in automation. *Hum Factors* 1995;**37**:381–94.
- Chopra V, Gesink BJ, de Jong J, et al. Does training on an anaesthesia simulator lead to improvement in performance? *Br J Anaesth* 1994;**73**:293–7.
- Gaba DM, Howard SK, Flanagan B, et al. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology* 1998;**89**:8–36.
- Marshall RL, Smith JS, Gorman PJ, et al. Use of a human patient simulator in the development of resident trauma management skills. *The Journal of Trauma: Injury, Infection, and Critical Care* 2001;**51**:17–21.
- Gaba DM. Two examples of how to evaluate the impact of new approaches to teaching. *Anesthesiology* 2002;**96**:1–2.
- Endsley MR. Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd annual meeting*. Santa Monica, CA: Human Factors Society, 1988:97–101.
- Endsley MR. Toward a theory of situation awareness in dynamic systems. *Hum Factors* 1995;**37**:32–64.
- Endsley MR. Theoretical underpinnings of situation awareness: a critical review. In: Endsley MR, Garland DJ, eds. *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000:3–32.
- Klein G. Analysis of situation awareness from critical incident reports. In: Endsley MR, Garland DJ, eds. *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000:51–71.
- Endsley MR. Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society 34th annual meeting*. Santa Monica, CA: Human Factors Society, 1990:41–5.
- Venturino M, Hamilton WL, Dvorchak SR. Performance-based measures of merit for tactical situation awareness. *Situational awareness in aerospace operations (AGARD-CP-478)*. Neuilly Sur Seine, France: NATO-AGARD, 1989, 4/1–4/5.
- Bell RH, Lyon DR. Using observer ratings to assess situation awareness. In: Endsley MR, Garland DJ, eds. *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000:129–46.
- Selcon SJ, Taylor RM, Koritsas E. Workload or situational awareness? TLX vs SART for aerospace systems design evaluation. *Proceedings of the Human Factors Society 36th annual meeting*. Santa Monica, CA: Human Factors Society, 1991:57–61.
- Collier SG, Folleso K. SACRI: a measure of situation awareness for nuclear power plant control rooms. In: Garland DJ, Endsley MR, eds. *Experimental analysis and measurement of situation awareness*. Daytona Beach, FL: Embry-Riddle University Press, 1995:115–22.
- Endsley MR. Situation awareness global assessment technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference (NAECON)*. New York: IEEE, 1988:789–95.
- Endsley MR, Mogford R, Allendoerfer K, et al. *Effect of free flight conditions on controller performance, workload, and situation awareness: a preliminary investigation of changes in locus of control using existing technology*. Atlantic City, NJ: Federal Aviation Administration William J Hughes Technical Center, 1997.
- Wickens CD. *Engineering psychology and human performance*. Glenview, Illinois: Scott Foresman and Company, 1984.
- Gaba DM. Human error in dynamic medical domains. In: Bogner MS, eds. *Human error in medicine*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994.
- Endsley MR. Designing for situation awareness in complex systems. *Proceedings of the Second International Workshop on Symbiosis of Humans, Artifacts and Environment*. Kyoto, Japan, 2001.
- Taylor RM. Situational awareness rating technique (SART): the development of a tool for aircrew systems design. *Situational awareness in aerospace operations*. Neuilly-Sur-Seine, France: NATO—AGARD, 1990, 3/1–3/17.
- Gaba DM, Howard SK, Small SD. Situation awareness in anaesthesiology. *Hum Factors* 1995;**37**:20–31.
- Endsley MR. Direct measurement of situation awareness: validity and use of SAGAT. In: Endsley MR, Garland DJ, eds. *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000:147–74.
- Sarter NB, Woods DD. Situation awareness: a critical but ill defined phenomenon. *Int J Aviat Psychol* 1991;**1**:45–57.
- Endsley MR. A methodology for the objective measurement of situation awareness. *Situational awareness in aerospace operations (AGARD-CP-478)*. Neuilly Sur Seine, France: NATO—AGARD, 1990, 1/1–1/9.
- Bolstad CA, Endsley MR. *Single versus dual scale range display investigation*. Hawthorne, CA: Northrop Corporation, 1990.
- Endsley MR. *Final report: situation awareness in an advanced strategic mission*. Hawthorne, CA: Northrop Corporation, 1989.
- Northrop Corporation. *Tactical simulation 2 test report: addendum 1 situation awareness test results*. Hawthorne, CA: Northrop Corporation, 1988.
- Hogg DN, Torralba B, Volden FS. *A situation awareness methodology for the evaluation of process control systems: studies of feasibility and the implication of use*. Storefjell, Norway: OECD Halden Reactor Project, 1993.
- Endsley MR, Sollenberger R, Nakata A, et al. *Situation awareness in air traffic control: enhanced displays for advanced operations*. Atlantic City, NJ: Federal Aviation Administration William J Hughes Technical Center, 1999.
- Endsley MR, Garland DJ, Shook RWC, et al. *Situation awareness problems in general aviation*. Marietta, GA: SA Technologies, Inc, 2000.
- Bolstad CA, Endsley MR, Howell C, et al. General aviation pilot training for situation awareness: an evaluation. *45th annual meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society, 2002.
- Strater LD, Endsley MR, Pleban RJ, et al. *Measures of platoon leader situation awareness in virtual decision making exercises*. Alexandria, VA: Army Research Institute, 2001.
- Strater LG, Reynolds JP, Faulkner LA, et al. PC-based tools to improve infantry situation awareness. *Human Factors and Ergonomics Society annual meeting*. Santa Monica, CA: Human Factors and Ergonomics Society, 2004.
- Cooke NJ. Cognitive task analysis for teams, CTA resource on-line seminar 2002. <http://www.ctaresource.com>.
- Cooke NJ, Salas E, Cannon-Bowers JA, et al. Measuring team knowledge. *Hum Factors* 2000;**42**:151–73.
- Murray DJ. Technical novelty or innovation in education. *Anesthesiology* 1998;**89**:1–2.