

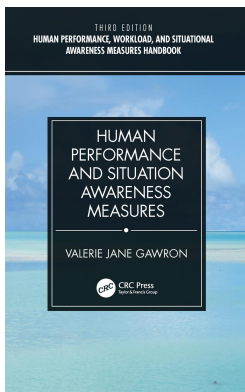
This article was downloaded by: 10.2.97.136

On: 09 Jun 2023

Access details: *subscription number*

Publisher: *CRC Press*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



Human Performance and Situation Awareness Measures

Gawron Valerie Jane

Measures of Situational Awareness

Publication details

<https://test.routledgehandbooks.com/doi/10.1201/9780429001024-3>

Gawron Valerie Jane

Published online on: 09 Jan 2019

How to cite :- Gawron Valerie Jane. 09 Jan 2019, *Measures of Situational Awareness from: Human Performance and Situation Awareness Measures* CRC Press

Accessed on: 09 Jun 2023

<https://test.routledgehandbooks.com/doi/10.1201/9780429001024-3>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

3

Measures of Situational Awareness

Situational Awareness (SA) is knowledge relevant to the task being performed. For example, pilots must know the state of their aircraft, the environment through which they are flying, and relationships between them, such as thunderstorms are associated with turbulence. It is a critical component of decision making and has been included in several models of decision making (e.g., Dorfel and Distelmaier model, 1997; see Figure 3.1). SA has three levels (Endsley, 1991): Level 1, perception of the elements in the environment; Level 2, comprehension of the current situation; and Level 3, projection of future status.

There are four types of SA measures: performance (also known as query methods, Durso and Gronlund, 1999), subjective ratings, simulation (also known as modeling, Golightly, 2015), and physiological measures. Individual descriptions of the first three types of measures of SA are provided in the following sections. Articles describing physiological measures of SA were written by French et al. (2003) and Vidulich et al. (1994). A flowchart to help select the most appropriate measure is given in Figure 3.2.

Note, another categorization of measures of SA is presented in Stanton et al. (2005). Their categories are: SA requirements analysis, freeze probe, real-time probe, self-rating probe, observer rating, and distributed SA. This team has also evaluated 17 SA measures for application in command, control, communication, computers, and intelligence (C4i) applications (Salmon et al., 2006). Their criteria were type of method SA requirements (analysis, freeze probe, real-time probe, self-rating, observer rating, performance measures, eye tracker), domain (air traffic control, civilian aviation, generic, military aviation, military infantry operations, nuclear power), team, subject matter experts required, training time, application time, tools needed, validation studies, advantages, and disadvantages. The authors concluded that all 17 measures were inadequate for the C4i application and recommended combining multiple methods such as performance measures, freeze probe, post-trial self-rating, and observer rating.

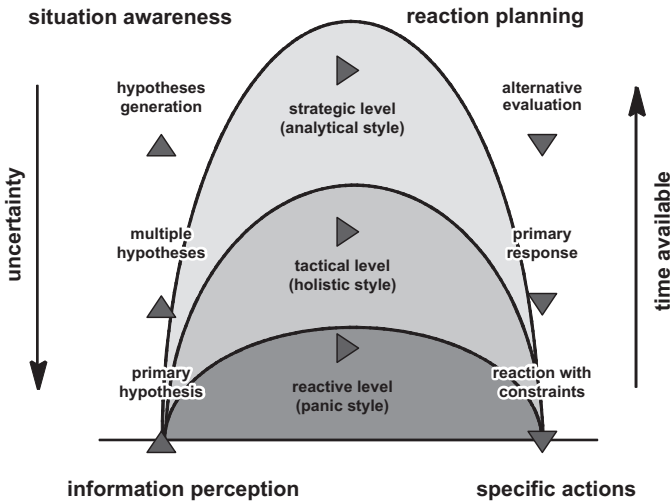


FIGURE 3.1
Decision making under uncertainty and time pressure (Dorfel and Distelmaier, 1997, p. 2).

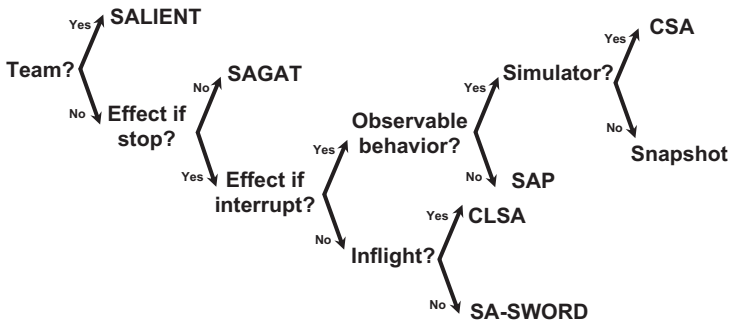


FIGURE 3.2
Guide to selecting a SA measure.

Sources

Dorfel, G., and Distelmaier, H. Enhancing situational awareness by knowledge-based user interfaces. Proceedings of the 2nd Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment, 197–205, 1997.

Durso, F.T., and Gronlund, S.D. Situation awareness. In F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumais, D.S. Lindsay, and M.T.H. Chi (Eds.) *Handbook of Applied Cognition* (pp. 283–314). New York: John Wiley and Sons, 1999.

- Endsley, M.R. Situation awareness in dynamic systems. In R.M. Taylor (Ed.) *Situational Awareness in Dynamic Systems (IAM Report 708)*. Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.
- French, H.T., Clark, E., Pomeroy, D., Clarke, C.R., and Seymour, M. Psycho-physiological measures of situation awareness. Proceedings of the Human Factors of Decision Making in Complex Systems, Part 6: Assessment and Measurement, Chapter 27, 291–298, August 2003.
- Golightly, D. Situation awareness. In J.R. Wilson and S. Sharples (Eds.) *Evaluation of Human Work* (pp. 515–548). Boca Raton: CRC Press, 2015.
- Salmon, P., Stanton, N., Walker, G., and Green, D. Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics* 37(2): 225–238, 2006.
- Stanton, N.A., Salmon, P.M., Walker, G.H., Barber, C., and Jenkins, D.P. *Human Factors Methods: A Practical Guide for Engineering and Design*. Gower ebook, December 2005.
- Vidulich, M.A., Stratton, M., Crabtree, M., and Wilson, G. Performance-based and physiological measures of situational awareness. *Aviation, Space, and Environmental Medicine* 65(5): A7–A12, 1994.

3.1 Performance Measures of SA

There are 12 performance measures of SA each described in a separate section below.

3.1.1 Cranfield Situation Awareness Scale (Cranfield-SAS)

General description – The Cranfield Situation Awareness Scale (Cranfield-SAS) was designed for flight instructors to rate students' SA in general aircraft handling, navigation, basic instrument flying, airways instrument flying, night flying, and commercial airline flight. It has two forms. The long version requires the instructor to rate from 1 (unacceptable) to 9 (excellent) student pilot knowledge (nine questions); understanding and anticipation of future events (five questions); management of stress, effort, and commitment (three questions); ability to attend, perceive, assimilate, and assess information (four questions); and overall awareness. The short form uses the same nine-point rating scale but uses only one rating per category (i.e., (1) knowledge; (2) understand and anticipate future events; (3) management; (4) capacity to perceive, attend, assimilate, and assess; and (5) overall awareness) (Dennehy, 1997).

Data requirements – Instructors must complete a rating for each question. Ratings can be made in flight or in a simulator and either during or after the flight.

Thresholds – 22 to 198 for the long form, 5 to 45 for the short form.

Source

Dennehy, K. *Cranfield – Situation Awareness Scale Users Manual (COA Report Number 9702)*. Bedford, England: Applied Psychology Unit Cranfield University, 1997.

3.1.2 Quantitative Analysis of Situation Awareness

General description – The Quantitative Analysis of Situation Awareness (QASA) is a freeze and query method developed by Edgar and Edgar (2007).

Strengths and limitations – Smith and Jamieson (2012) reported a significant difference in QASA scores between manual and information analysis groups in a Cabin Air Management System simulation.

Data requirements – Develop relevant probe questions.

Thresholds – 0 to total number of probe questions.

Sources

Edgar, G.K., and Edgar, H.E. Using signal detection theory to measure situation awareness. The technique, the tool, the test, the way forward. In J. Noyes and M. Cook (Eds.) *Decision Making in Complex Environments* (pp. 373–385). Aldershot, UK: Ashgate, 2007.

Smith, A.G., and Jamieson, G.A. Level of automation effects on situation awareness and functional specificity on automation reliance. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, 2113–2117, 2012.

3.1.3 Quantitative Analysis of Situational Awareness (QUASA)

General description – The Quantitative Analysis of Situational Awareness (QUASA) combines the accuracy of true/false situational awareness probe questions with the self-rating of the confidence of the answer (very high, high, moderate, low, very low).

Strengths and limitations – McGinness (2004) reported that there were significant differences in SA among five teams performing command and control exercises.

Data requirements – A plot of the participant's confidence against the proportion correct of the probe questions is a calibration curve. SA accuracy is measured as the proportion correct. Perceived accuracy is derived from the five-point rating scale.

Thresholds – 0 to 100 for proportion correct, 1 to 5 for confidence.

Source

McGinness, B. Quantitative Analysis of Situational Awareness (QUASA): Applying signal detection theory to true/false probes and self-ratings. 9th International Command and Control Research and Technology Symposium, 159–178, 2004.

3.1.4 SA ANALYSIS TOOL (SAVANT)

General description – The SA ANALYSIS Tool (SAVANT) uses data from the Air Traffic Control operational systems to probe SA. SAVANT presents a query for three seconds. Queries are aircraft-pair and sector-based questions.

Strengths and limitations – Willems and Heiney (2002) reported significant differences in the time to respond correctly to SAVANT questions. Specifically, Air Traffic Controllers responded faster on the radar side than on the data side. Further, they responded faster to future SA queries when the task load was low. Presentation of the SAVANT screen causes a blind period for the radar or data screen.

Data requirements – Air Traffic Control Subject Matter Experts are required to complete the SAVANT forms with correct answers.

Thresholds – Time to correct responses have a minimum of 0 ms.

Source

Willems, B., and Heiney, M. *Decision Support Automation Research in the En Route Air Traffic Control Environment (DOT/FAA/CT-TN01/10)*. Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center, January 2002.

3.1.5 SALSA

General description – The acronym is German for “Measuring Situation Awareness of Area Controllers within the Context of Automation” (Haus and Eyferth, 2003, p. 422). It combines responses to questions posed at simulation freeze with expert ratings of the relevance of the probe question, uses cued recall to reduce confusion, and limits the probe questions to one aircraft at a single freeze.

Strengths and limitations – Haus and Eyferth (2003) measured SALSA for 11 air traffic controllers over 45 minutes of simulated traffic. They concluded that SALSA was feasible but recommended the development of a taxonomy of air traffic characteristics.

Data requirements – Probe questions, ability to freeze a simulation, and availability of experts to provide ratings of relevance.

Thresholds – Not stated.

Source

Hauss, Y., and Eyferth, K. Securing future ATM-concepts' safety by measuring situation awareness in ATC. *Aerospace Science and Technology* 7(6): 417–427, 2003.

3.1.6 Shared Awareness Questionnaire

General description – The Shared Awareness Questionnaire requires participants to answer objective questions related to an experiment or exercise. Examples include locations critical to mission success and where these locations were on a map (Prytz et al., 2015). The questionnaire is scored on inter-rater agreement and accuracy.

Strengths and limitations – The Shared Awareness Questionnaire was used in an emergency management exercise (Prytz et al., 2015). Fifty participants completed the questionnaire after the exercise was concluded.

Thresholds – Zero to 100% inter-rater agreement and 0 to 100% accuracy.

Source

Prytz, E., Rybing, J., Jonson, C., Petterson, A., Berggren, P., and Johansson, B. An exploratory study of low-level shared awareness measure using mission-critical locations during an emergency response. Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, 1152–1156, 2015.

3.1.7 Situational Awareness Global Assessment Technique (SAGAT)

General description – Among the most well-known measure of SA is the Situational Awareness Global Assessment Technique (SAGAT) (Endsley, 1988b). SAGAT was designed around real-time, human-in-the-loop simulation of a military cockpit but could be generalized to other systems. SAGAT uses a graphical computer program for the rapid presentation of queries and data collection. Using SAGAT, the simulation is stopped at random times and the operators are asked questions to determine their SA at that particular point in time. Participants' answers are compared with the correct answers that have been simultaneously collected in the computer database.

“The comparison of the real and perceived situation provides an objective measure of SA” (Endsley, 1988a, p. 101).

Strengths and limitations – The SAGAT technique has been tested in several studies, which demonstrated: (1) empirical validity (Endsley, 1989, 1990b) – the technique of freezing the simulation did not impact participant performance and participants were able to reliably report SA knowledge for up to six minutes after a freeze without memory decay problems; (2) predictive validity (Endsley, 1990b) – linking SAGAT scores to participant performance; and (3) content validity (Endsley, 1990a) – showing appropriateness of the queries used (for an air-to-air fighter cockpit).

Bolstad (1991) correlated SAGAT with scores from potential SA predictor tests. The participants were Northrop employees with tactical flight experience. The top correlations with SAGAT were: attention sharing tracking level (+0.717), immediate/delayed memory total errors (+0.547), encoding speed (+0.547), dot estimation (+0.415), and Group Embedded Figures Test number correct (+0.385).

Bolstad and Endsley (2003) used Goal-Directed Cognitive Task Analysis to develop SAGAT queries. Jones et al. (2004) extended this work to build a toolkit (Designer’s Situation Awareness Toolkit) to help develop SAGAT queries.

SAGAT provides objective measures of SA across all of the operators’ SA requirements that can be computed in terms of errors or percent correct and can be treated accordingly. However, Sarter and Woods (1991) suggest that SAGAT does not measure SA but rather measures what pilots can recall. Further, Fracker and Vidulich (1991) identified two major problems with the use of explicit measures of SA, such as SAGAT: (1) decay of information and (2) inaccurate beliefs.

Other researchers have simply asked participants to indicate locations of objects on maps. This technique is sometimes referred to as mini-SAGAT or snapshots. It has been used in a wide variety of military and civilian tasks. For example, firefighters have been asked to mark the location of a fire on a map. SA is measured as the deviation between the actual fire and the indicated location on the map (Artman, 1999). Another example is recall of the positions of traffic in front of a participant’s vehicle (Johannsdottir and Herdman, 2010). Still another example was used to assess SA of unmanned system operators to locate targets (McDermott and Fisher, 2013). Tippey et al. (2017) used reaction time (RT) to situation awareness probes to evaluate vibrotactile and embedded text cues in general aviation weather alerting.

Vidulich et al. (1995) used a SAGAT-like memory probe to measure SA on a PC-based flight simulator performance. The SAGAT-like measure was sensitive to differences in information presented during the simulation but not to difficulty, thus enabling the differentiation of SA and workload.

SAGAT has been used in many domains including: aviation, command and control, driving, energy, health care, robotics, and submarine display design.

Aviation. Many of the early uses of SAGAT were in military aviation. Fracker (1989) stated from a simulated air threat study “the present data encourage the use of memory probes to measure situation awareness.” Fracker (1991) evaluated SA measures in a simulated combat task. Test-retest correlations were significant for identity accuracy, identity latency, envelope sensitivity, and kill probability but not for location error and avoidance failure. Only identity latency was significantly correlated with kill probability. This correlation was used to assess criterion validity. Construct validity was evaluated from correlations of SA metrics with avoidance failure. Three correlations were significant: (1) identity latency, (2) envelope sensitivity, and (3) kill probability. Two correlations were not significant: (1) location error and (2) identity accuracy. There were three significant correlations among SA metrics: (1) identity accuracy and location error, (2) identity accuracy and identity latency, and (3) envelope sensitivity and latency. A 0.10 alpha was used to determine significance. Three correlations were not significant: (1) identity latency and location error, (2) envelope sensitivity and location error, and (3) envelope sensitivity and identify accuracy.

Crooks et al. (2001) asked 165 undergraduate students to rate their SA using SAGAT for a simulated aircraft identification task. The results were mixed. There were significant differences in SAGAT for corridor, size, and direction but not for speed, range, angle, and Identify Friend or Foe. Further, for range and angle, there was a significant difference in the direction opposite the direction hypothesized.

Endsley (1995) did not find any significant difference in SAGAT scores as a function of time (0 to 400 s) between event and response nor did she see any performance decrements in a piloting task between stopping the simulation and not stopping the simulation. Bolstad et al. (2002, 2003) used SAGAT to measure the effectiveness of time-sharing training on enhancing SA. The participants were 24 pilots. The data were collected using the Microsoft *Flight Sim 2000*. There was a significant increase in SA after the training. Those with training went from 22% correct awareness of wind direction prior to training to 55% after training. Those without training went from 58% pre-flight to 36% postflight.

Snow and Reising (2000) found no significant correlation between SAGAT and SA-SWORD and further only SA-SWORD showed statistically significant effects of visibility and synthetic terrain type in a simulated flight.

SAGAT has also been used in commercial aviation. Boehm-Davis et al. (2010) found no significant differences in awareness for pilots in using Data Comm than using voice. The participants were 24 transport pilots. The data were collected in a simulator.

Prince et al. (2007) used a SAGAT-like approach to measure team situation awareness. In a high-fidelity flight simulator, team situation awareness was measured by two instructors who had completed a three-day training course. The mean correlation between the two instructors on the 48 situation awareness items was $r = + 0.88$. There were also significant correlations

between team situation awareness and ratings of problem solving for icing, boost pump, and fire emergencies. In a low-fidelity simulator, team simulation awareness was measured by responses to flight status. Questions included: current altitude and heading; controlling agency; last air traffic control (ATC) call; weather at destination, emergency airport, airport bearing and range; and distance to, clock position of, and altitude of traffic. The data were from 41 military aircrews. There were significant correlations between responses to flight status questions and ratings of problem solving for boost pump and fire emergencies but not for icing.

Vu et al. (2010) reported both percent error and probe latency differences for pilots applying different concepts of operation.

SAGAT has been used in evaluations of Air Traffic Control systems as well. Endsley and Rodgers (1996) evaluated SAGAT scores associated with 15 Air Traffic Scenarios associated with operational errors. Their participants were 20 Air Traffic Control Specialists. Low percentage correct on recall items was associated with assigned clearances complete (23.2%), speed (28%), turn (35.1%), and call sign numeric (38.4%). Endsley et al. (2000) compared SAGAT, on-line probes, Situational Awareness Rating Technique (SART), and observer ratings during an Air Traffic Control simulation using 10 experienced air traffic controllers. Only SAGAT Level 2 SA queries were associated with significant differences between display conditions under test.

In a study in the same domain, Jones and Endsley (2004) reported a significant correlation between SAGAT and real-times probes. Their participants were Air Traffic Controllers. Also in Air Traffic Control, Kaber et al. (2006) reported significant differences in level 1 SA between automation modes with information acquisition resulting in the highest SA and action implementation in the lowest SA.

Sethumandhavan (2011a) used SAGAT to measure SA during a simulated air traffic control task using one of four levels of automation (information acquisition, information analysis, decision and action selection, and action implementation automation) and before and after an automation failure. There was significantly higher SA in the information analysis automation condition than the other three automation conditions. Further, participants had significantly higher SA before the occurrence of an automation failure than after it. This author also reported (2011b) that meta-SA (confidence of a participant's own ability to recall aircraft attributes during an air traffic control simulation) was significantly related to SAGAT scores.

In an Air Traffic Control task, Strybel et al. (2009) reported that accuracy was significantly lower for multiple choice probes than for yes/no or rating probes. Willems and Heiney (2002) used SAGAT to evaluate air traffic control SA to airspace sector-based queries. They reported that correct responses for aircraft placement was only about 20%. Further, there was a decrease in SA as task load increased and was less on the data side than on the radar side. Durso et al. (1999) also reported low percent correct scores for Air Traffic Controllers.

In an unusual UAV study, Draper et al. (2000) reported that haptic cues significantly increased the accuracy of identifying turbulence wind direction for UAV pilots. Fern and Shively (2011) also evaluating UAV operator SA used percent correct responses to SA probe questions. They reported significant differences in SA among some but not all display types.

Command and Control. In a command and control domain experiment, Jones and Endsley (2000) reported that response accuracy to real-time probes was associated with significant differences between scenarios (war and peace) but response time did not. For SAGAT, 11 or 21 queries were significantly different between the two scenarios but the aggregate SAGAT score was not. Overall SART scores were significantly different between the two scenarios. French and Hutchinson (2002) applied SAGAT in a military ground force command and control exercise. There was an increase in situation awareness across time. The authors recommended a broader range of probes for future research. Hall et al. (2010) used latency to Commander's Critical Information Report (CCIR) to compare a baseline and a prototype command and control display. The prototype was associated with significantly shorter response latencies.

Lampton et al. (2006) compared responses to event and location to a SA Measure of Team Communications (SAMTC) derived from ratings of team communication on seeking updates, identifying problems, and preventing errors and an adaptation of the Situation Awareness Behaviorally Anchored Rating Scale (SABARS) to rate from 1 to 10 acquisition and dissemination of SA. The authors concluded that all three SA measures were able to measure SA.

In a review of three experiments (military personnel recovery, military command and control, and combat vehicle maneuvering), Cuevas and Bolstad (2010) evaluated the correlation of the team leader's SA using SAGAT and the team members' SA also using SAGAT. There were significant positive correlations for the military personnel recovery and the combat vehicle maneuvering. The authors suggested that the lack of a significant correlation in the command and control experiment was due to missing data. In a separate analysis of the recovery task, Saner et al. (2010) reported significant predictions of shared SA from experience similarity, workload similarity, organizational hub distance, and communication distance but not for shared knowledge.

Parush and Ma (2012) used the mean number of correct responses to the true/false situation awareness statements to assess a team display for managing forest fire incidents. In the following year, Parush and Rustanjaja (2013) used the same method to assess SA for firefighters working in apartments and office buildings.

Giacobe (2013) used SAGAT to compare SA of novice and experienced cyber analysts using text or graphic information. He also measured workload using NASA TLX and perceived awareness using SART. There were no significant effects of either experience or interface style on SAGAT score.

Driving. Bolstad (2000) used queries related to driving to assess differences in SA ability with driver age.

Energy. Hogg et al. (1995) adapted SAGAT to develop the Situation Awareness Control Room Inventory (SACRI) to evaluate nuclear power plant operator's SA. They completed four simulator studies using SACRI and concluded that it was useful in design assessments when used with time and accuracy measures. Similarly, Lenox et al. (2011) adapted SAGAT for electric power transmission and distribution operations.

Tharanathan et al. (2010) evaluated a processing monitoring display. Level 1 SA was measured as the number of process changes identified by the operator, level 1 by the accuracy of responses to probe questions, and level 3 by the number of warning flags predicted. There were significant differences between displays and scenario complexity for SA levels 1 and 2 measures but only a significant difference for scenario complexity on the level 3 SA measure.

In an evaluation of process control displays, Bowden and Rusnock (2015) reported only one significant effect on SAGAT, specifically, graphic displays were associated with higher SA than numeric displays. There was no significant difference between functionally grouped and spatially mapped information displays.

Health care. Luz et al. (2010) used SAGAT ratings after a simulated Mastoidectomy to evaluate the effectiveness of image-guided navigation during surgery. They reported no significant differences in ratings between completing the procedure manually than with the imagery system. Saetrenik (2012) used freeze probes in an emergency handling training scenario. There was a significant difference in SA (average of the nine probes) across the five teams but no significant difference in SA by the role each team member played.

Robotics. Riley and Strater (2006) used SAGAT to compare SA in four robot control modes and reported no significant differences.

Submarine Display Design. Huf and French (2004) reported SART indicating low understanding of the situation in a Virtual Submarine Environment while SAGAT indicated high understanding as evidences by correct responses. Loft et al. (2015) compared SPAM, SAGAT, ATWIT, NASA TLX, and SART ratings from 117 undergraduates performing three submarine tasks: contact classification, closest point of approach, and emergency surface. SPAM did not significantly correlate with SART but did with ATWIT and NASA TLX.

Data requirements – The proper queries must be identified prior to the start of the experiment. Landry and Yoo (2012), based on a Monte Carlo simulation of responses to SA queries, reported that statistical power reduces from 0.8 for 50 queries to 0.58 for 30 queries, 0.24 for 10 queries, and 0.15 for five queries. Endsley (2000) stated that a detailed analysis of SA requirements is required to develop the queries.

Thresholds – Tolerance limits for acceptable deviance of perceptions from real values on each parameter should be identified prior to the start of the experiment.

Sources

- Artman, H. Situation awareness and co-operation within and between hierarchical units in dynamic decision making. *Ergonomics* 42(11): 1404–1417, 1999.
- Boehm-Davis, D.A., Gee, S.K., Baker, K., and Medina-Mora, M. Effect of party line loss and delivery format on crew performance and workload. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 126–130, 2010.
- Bolstad, C.A. Age-related factors effecting the perception of essential information during risky driving situations. In D.B. Kaber and M.R. Endsley (Eds.) Proceedings of the First Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, October 15–19, 2000.
- Bolstad, C.A. Individual pilot differences related to situation awareness. Proceedings of the Human Factors Society 35th Annual Meeting, 52–56, 1991.
- Bolstad, C.A., and Endsley, M.R. Measuring shared and team situation awareness in the Army's future objective force. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 369–373, 2003.
- Bolstad, C.A., Endsley, M.R., Howell, C.D., and Costello, A.M. General aviation pilot training for situation awareness: An evaluation. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, 21–25, 2002.
- Bolstad, C.A., Endsley, M.R., Howell, C.D., and Costello, A.M. The effect of time-sharing training on pilot situation awareness. Proceedings of the 12th International Symposium on Aviation Psychology, 140–145, 2003.
- Bowden, J.R., and Rusnock, C.F. Evaluation of human machine interface design factors on situation awareness and task performance. Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, 1361–1365, 2015.
- Crooks, C.L., Hu, C.Y., and Mahan, R.P. Cue utilization and situation awareness during simulated experience. Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, 1563–1567, 2001.
- Cuevas, H.M., and Bolstad, C.A. Influence of team leaders' situational awareness on their team's situational awareness and performance. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 309–313, 2010.
- Draper, M.H., Ruff, H.A., Repperger, D.W., and Lu, L.G. Multi-sensory interface concepts supporting turbulence detection by UAV controllers. Proceedings of the First Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 107–112, 2000.
- Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D., and Manning, C.A. *Situation Awareness as a Predictor of Performance in En Route Air Traffic Controllers (DOT/FAA/AM-99/3)*. Washington, DC: Office of Aviation Medicine, January 1999.

- Endsley, M.R. Design and evaluation for situation awareness enhancement. Proceedings of the 32nd Annual Meeting of the Human Factors Society, 97–101, 1988a.
- Endsley, M.R. Situational awareness global assessment technique (SAGAT). Proceedings of the National Aerospace and Electronics Conference, 789–795, 1988b.
- Endsley, M.R. A methodology for the objective measurement of pilot situation awareness. Presented at the AGARD Symposium on Situation Awareness in Aerospace Operations, Copenhagen, Denmark, October 1989.
- Endsley, M.R. *Situation Awareness in Dynamic Human Decision Making: Theory and Measurement (NORDOC 90-49)*. Hawthorne, CA: Northrop Corporation, 1990a.
- Endsley, M.R. Predictive utility of an objective measure of situation awareness. Proceedings of the Human Factors Society 34th Annual Meeting, 41–45, 1990b.
- Endsley, M.R. Toward a theory of situational awareness in dynamic systems. *Human Factors* 37(1): 32–64, 1995.
- Endsley, M.R. Direct measurement of situation awareness: Validity and use of SAGAT. In M.R. Endsley and D.J. Garland (Eds.) *Situation Awareness: Analysis and Measurement* (pp. 131–157). Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- Endsley, M.R., and Rodgers, M.D. Attention distribution and situation awareness in air traffic control. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, 82–85, 1996.
- Endsley, M.R., Sollenberger, R., and Stein, S. Situation awareness: A comparison of measures. Proceedings of the First Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, October 15–19, 2000.
- Fern, L., and Shively, J. Designing airspace displays to support rapid immersions for UAS handoffs. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 81–85, 2011.
- Fracker, M.L. Attention allocation in situation awareness. Proceedings of the Human Factors Society 33rd Annual Meeting, 1396–1400, 1989.
- Fracker, M.L. *Measures of Situation Awareness: An Experimental Evaluation (AL-TR-1991-0127)*. OH: Wright-Patterson Air Force Base, October 1991.
- Fracker, M.L., and Vidulich, M.A. Measurement of situation awareness: A brief review. In R.M. Taylor (Ed.) *Situational Awareness in Dynamic Systems (IAM Report 708)* (pp. 795–797). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.
- French, H.T., and Hutchinson, A. Measurement of situation awareness in a C4ISR experiment. Proceedings of the 7th International Command and Control Research and Technology Symposium. CCRP, Washington, DC, 2002.
- Giacobe, N.A. A picture is worth a thousand words. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 172–176, 2013.
- Hall, D.S., Shattuck, L.G., and Bennett, K.B. Evaluation of an ecological interface designed for military command and control. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 423–427, 2010.
- Hogg, D.N., Folleso, K., Strand-Volden, F., and Torralba, B. Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics* 38(11): 2394–2413, 1995.

- Huf, S., and French, H.T. Situation awareness in a networked virtual submarine. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 663–667, 2004.
- Johannsdottir, K.R., and Herdman, C.M. The role of working memory in supporting drivers' situation awareness for surrounding traffic. *Human Factors* 52(6): 663–673, 2010.
- Jones, D.G., and Endsley, M.R. Can real-time probes provide a valid measure of situation awareness? Proceedings of the First Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 245–250, 2000.
- Jones, D.G., and Endsley, M.R. Use of real-time probes for measuring situation awareness. *The International Journal of Aviation Psychology* 14(4): 343–367, 2004.
- Jones, D.G., Endsley, M.R., Bolstad, M., and Estes, G. The designer's situation awareness toolkit: Support for user-centered design. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 653–657, 2004.
- Kaber, D.B., Perry, C.M., Segall, N., McClernon, C.K., and Prinzl, L.J. Situation awareness implication of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics* 36(5): 447–462, 2006.
- Lampton, D.R., Riley, J.M., Kaber, D.B., Sheik-Nainar, M.A., and Endsley, M.R. Use of immersive virtual environments for measuring and training situation awareness. Presented at the U.S. Army Science Conference Orlando, Florida, November 2006.
- Landry, S.J., and Yoo, H. Sampling error and other statistical problems with query-based situation awareness measures. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 292–296, 2012.
- Lenox, M.M., Connors, E.S., and Endsley, M.R. A baseline evaluation of situation awareness for electric power system operation supervisors. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 2044–2048, 2011.
- Loft, S., Bowden, V., Braithwaite, J., Morrell, D.B., Huf, S., and Durso, F.T. Situation awareness measures for simulated submarine track management. *Human Factors* 57(2): 298–310, 2015.
- Luz, M., Mueller, S., Strauss, G., Dietz, A., Meixenberger, J., and Manzey, D. Automation in surgery: The impact of navigation-control assistance on the performance, workload and situation awareness of surgeons. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 889–893, 2010.
- McDermott, P.L., and Fisher, A. Methodologies for assessing situation awareness of unmanned system operators. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 167–171, 2013.
- Parush, A., and Ma, C. Team displays work, particularly with communication breakdown: Performance and situation awareness in a simulated forest fire. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 383–387, 2012.
- Parush, A., and Rustanjaja, N. The impact of sudden events and spatial configuration on the benefits of prior information to situation awareness and performance. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 1395–1399, 2013.

- Prince, C., Ellis, J.E., Brannick, M.T., and Salas, E. Measurement of team situation awareness in low experience level aviators. *International Journal of Aviation Psychology* 17(1): 41–57, 2007.
- Riley, J.M., and Strater, L.D. Effects of robot control mode on situational awareness and performance in a navigation task. Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, 540–544, 2006.
- Saetrenik, B. A controlled field study of situation awareness measures and heart rate variability in emergency handling teams. Proceedings of the eHuman Factors and Ergonomics Society 56th Annual Meeting, 2006–2010, 2012.
- Saner, L.D., Bolstad, C.A., Gonzalez, C., and Cuevas, H.M. Predicting shared situational awareness in teams; A case of differential SA requirements. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 314–318, 2010.
- Sarter, N.B., and Woods, D.D. Situational awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology* 1(1): 45–57, 1991.
- Sethumandhavan, A. Effects of first automation failure on situation awareness and performance in an air traffic control task. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 350–354, 2011a.
- Sethumandhavan, A. Knowing what you know: The role of meta-situation awareness in predicting situation awareness. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 360–364, 2011b.
- Snow, M.P., and Reising, J.M. Comparison of two situation awareness metrics: SAGAT and SA-SWORD. Proceedings of the IEA 2000/HFES 2000 Congress, vol. 3, pp. 49–52, 2000.
- Strybel, T.Z., Minakata, K., Nguyen, J., Pierce, R., and Vu, K.L. Optimizing online situation awareness probes in air traffic management tasks. In M.J. Smith and G. Salvendy (Eds.) *Human Interface, Part II, HCI 2009, LNCS 5618* (pp. 845–854). Berlin: Springer, 2009.
- Tharanathan, A., Laberge, J., Bullemer, P., Reising, D.V., and McLain, R. Functional versus schematic overview displays: Impact on operator situation awareness in process monitoring. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 319–323, 2010.
- Tippey, K.G., Roady, T., Rodriguez-Paras, C., Brown, L.J., Rantz, W.G., and Ferris, T.K. General aviation weather alerting: The effectiveness of different visual and tactile display characteristics in supporting weather-related decision making. *The International Journal of Aerospace Psychology* 27(3–4): 121–136, 2017.
- Vidulich, M.A., McCoy, A.L., and Crabtree, M.S. The effect of a situation display on memory probe and subjective situational awareness metrics. Proceedings of the 8th International Symposium on Aviation Psychology, 765–768, 1995.
- Vu, K.L., Strybel, T.Z., Kraut, J., Bacon, P., Minakata, K., Nguyen, J., Rottermann, A., Battiste, V., and Johnson, W. Pilot and controller workload and situation awareness with three traffic management concepts (Paper 978-1-4244-6618-4). IEEE 29th Digital Avionics Systems Conference, 4.A.5-1–4.A.5-10, 2010.
- Willems, B., and Heiney, M. *Decision Support Automation Research in the En Route Air Traffic Control Environment (DOT/FAA/CT-TN01/10)*. Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center, January 2002.

3.1.8 Situational Awareness Linked Instances Adapted to Novel Tasks

General description – The Situational Awareness Linked Instances Adapted to Novel Tasks (SALIENT) was developed to measure team SA. The SALIENT methodology requires five phases: (1) identify team SA behaviors (see Table 3.1), (2) develop scenarios, (3) define acceptable responses, (4) write a script, and (5) create a structured form with columns for scenarios and responses.

Strengths and limitations – SALIENT has been validated using twenty undergraduate students in a four-hour tabletop helicopter simulation. Inter-rater reliability was $r = +0.94$. There were significant correlations between SALIENT score and communication frequency ($r = +0.74$), between SALIENT score and performance ($r = +0.63$). There were no significant correlations between SALIENT score and the teams' shared mental model ($r = -0.04$). Additional validation data are available in Muniz et al. (1998a) and Bowers

TABLE 3.1

Generic Behavioral Indicators of Team SA (Muniz et al., 1998b)

Demonstrated Awareness of Surrounding Environment

Monitored environment for changes, trends, abnormal conditions
Demonstrated awareness of where he/she was

Recognized Problems

Reported problems
Located potential sources of problem
Demonstrated knowledge of problem consequences
Resolved discrepancies
Noted deviations

Anticipated a Need for Action

Recognized a need for action
Anticipated consequences of actions and decisions
Informed others of actions taken
Monitored actions

Demonstrated Knowledge of Tasks

Demonstrated knowledge of tasks
Exhibited skill time sharing attention among tasks
Monitored workload
Shared workload within station
Answered questions promptly

Demonstrated Awareness of Information

Communicated important information
Confirmed information when possible
Challenged information when doubtful
Re-checked old information
Provided information in advance
Obtained information of what is happening
Demonstrated understanding of complex relationship
Briefed status frequently

et al. (1998). In a large experiment (80 men and 180 women), Fink and Major (2000) compared SART and SALIENT related to a helicopter flight simulation game (Werewolf v Comanche). The authors reported that SALIENT had the better measurement characteristics but SART enabled them to assess the interaction between the game and the operator. Kardos (2003) applied the SALIENT methodology to develop a checklist of behavioral SA indicators during development of Tactics, Techniques, and Procedures (TTPs) for the Australian Defence Science and Technology Organisation (DSTO). She described the difficulty in observing some of the behaviors identified such as “monitor others.” She recommended excluding some of the behaviors due to lack of applicability, constancy, and not being observed.

Data requirements – Although the generic behaviors in Table 3.1 can be used, scenarios, responses, scripts, and report forms must be developed for each team task.

Thresholds – Not stated.

Sources

- Bowers, C., Weaver, J., Barnett, J., and Stout, R. Empirical validation of the SALIENT methodology. Proceedings of the North Atlantic Treaty Organization Research and Technology Organization Meeting 4 (RTO-MP-4, AC/23(HFM)TP/2), 12-1-12-6, April 1998.
- Fink, A.A., and Major, D.A. Measuring situation awareness: A comparison of three techniques. Proceedings of the 1st Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 2000.
- Kardos, M. *Behavioral Situation Awareness Measures and the Use of Decision Support Tools in Exercise Prowling Pegasus*. Edinburgh, South Australia: DSTO Systems Sciences Laboratory, 2003.
- Muniz, E.J., Salas, E., Stout, R.J., and Bowers, C.A. The validation of a team situational awareness measure. Proceedings for the 3rd Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment. Naval Air Warfare Center Aircraft Division, Patuxent River, MD, 183-190, 1998a.
- Muniz, E.J., Stout, R.J., Bowers, C.A., and Salas, E. A methodology for measuring team situational awareness: Situational Awareness Linked Indicators Adapted to Novel Tasks (SALIENT). Proceedings of the NATO Human Factors and Medicine Panel on Collaborative Crew Performance in Complex Systems, 20-24, 1998b.

3.1.9 Situation Present Assessment Method (SPAM)

General description – The Situation Present Assessment Method (SPAM) uses response latency to queries as the measure of SA. The queries are asked when

all information necessary to answer the query is present, hence eliminating a memory component. It was developed by Durso et al. (1995).

Strengths and limitations – Strybel et al. (2016) used SPAM to compare the situation awareness of retired Air Traffic Controllers using four different separation assurance and spacing concepts in enroute and transitional sectors. The authors reported that in the enroute sector SA was highest when controllers managed separation assurance and the automation managed the spacing. In the transitional sector, SA was highest when the controllers managed both functions.

Vu et al. (2009) reported students training in Air Traffic Control answered significantly more probe questions than retired Air Traffic Controllers.

Durso et al. (1999) applied SPAM to the comparison of Air Traffic Control scenarios. There were no significant differences. Baker et al. (2012) reported no significant difference in SPAM scores between data comm and voice communication for pilots using a low-fidelity flight simulator. Pierce (2012) reported that SPAM reduced the number of correctly handled aircraft.

In a domain other than Air Traffic Control, Schuster et al. (2012) reported significantly higher SPAM scores (i.e., number of questions answered correctly) for mission 2 than mission 3 with a human-robot team task of identifying soldiers presented in a simulated Military Operations in Urban Terrain (MOUT).

Loft et al. (2015b) reported no significant differences in SPAM accuracy or time to respond to SPAM queries between two levels of task load. Loft et al. (2015a) compared SPAM, SAGAT, ATWIT, NASA TLX, and SART ratings from 117 undergraduates performing three submarine tasks: contact classification, closest point of approach, and emergency surface. SPAM did not significantly correlate with SART but did with ATWIT and NASA TLX.

Data requirements – Questions must be identified as well as start time for measuring response latency.

Thresholds – Not stated.

Sources

- Baker, K.M., DiMare, S.K., Nelson, E.T., and Boehm-Davis, D.A. Effect of data communications on pilot situation awareness, decision making, and workload. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 1787–1788, 2012.
- Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D., and Manning, C.A. *Situation Awareness as a Predictor of Performance in En Route Air Traffic Controllers (DOT/FAA/AM-99/3)*. Washington, DC: Office of Aviation Medicine, January 1999.

- Durso, F.T., Truitt, T.R., Hackworth, C.A., Crutchfield, J.M., Nikolic, D., Moertl, P.M., Ohrt, D., and Manning, C.A. Expertise and chess: Comparing situation awareness methodologies. Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness, 295–303, 1995.
- Loft, S., Bowden, V., Braithwaite, J., Morrell, D.B., Huf, S., and Durso, F.T. Situation awareness measures for simulated submarine track management. *Human Factors* 57(2): 298–310, 2015a.
- Loft, S., Sadler, A., Braithwaite, J., and Huff, S. The chronic detrimental impact of interruptions in a simulated submarine track management task. *Human Factors* 57(8): 1417–1426, 2015b.
- Pierce, R.S. The effect of SPAM administration during a dynamic simulation. *Human Factors* 54(5): 838–848, 2012.
- Schuster, D., Keebler, J.R., Jentsch, F., and Zuniga, J. Comparison of SA measurement techniques in a human-robot team task. Proceedings of the Human Factors Society 56th Annual Meeting, 1713–1717, 2012.
- Strybel, T.Z., Vu, K.L., Chiappe, D.L., Morgan, C.A., Morales, G., and Battiste, V. Effects of NextGen concepts of operation for separation assurance and interval management on air traffic controller situation awareness, workload, and performance. *International Journal of Aviation Psychology* 26(1–2): 1–14, 2016.
- Vu, K.P.L., Minakata, K., Nguyen, J., Kraut, J., Raza, H., Battiste, V., and Strybel, T.Z. Situation awareness and performance of student versus experienced air traffic controllers. In M.J. Smith and G. Salvendy (Eds.) *Human Interface, Part II, HCII 2009, LNCS5618* (pp. 865–874), 2009.

3.1.10 Tactical Rating of Awareness for Combat Environments (TRACE)

General description – The Tactical Rating for Awareness for Combat Environments (TRACE) estimates SA from the accuracy of responses to periodic situation reports (SITREP) and the elapsed time from the request for the SITREP and the completion of all SITREP line items (Hall, 2009).

Strengths and limitations – Hall et al. (2010) used TRACE to compare a baseline and a prototype command and control display. The prototype was associated with both significantly shorter response latencies and significantly greater number of correct responses.

Data requirements – Questions must be identified as well as start time for measuring response time.

Thresholds – Not stated.

Sources

- Hall, D.S. Raptor: An empirical evaluation of an ecological interface designed to increase warfighter cognitive performance. Master's Thesis. Naval Postgraduate School, Monterey, CA, June 2009.

Hall, D.S., Shattuck, L.G., and Bennett, K.B. Evaluation of an ecological interface designed for military command and control. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 423–427, 2010.

3.1.11 Temporal Awareness

General description – Temporal awareness has been defined as “the ability of the operator to build a representation of the situation including the recent past and the near future” (Grosjean and Terrier, 1999, p. 1443). It has been hypothesized to be critical to process management tasks.

Strengths and limitations – Temporal awareness has been measured as the number of temporal and ordering errors in a production line task, number of periods in which temporal constraints were adhered to, and the temporal landmarks reported by the operator to perform his or her task. Temporal landmarks include relative ordering of production lines and clock and mental representation of the position of production lines.

Data requirements – Correct time and order must be defined with tolerances for error data. Temporal landmarks must be identified during task debriefs.

Thresholds – Not stated.

Source

Grosjean, V., and Terrier, P. Temporal awareness: Pivotal in performance? *Ergonomics* 42(11): 1443–1456, 1999.

3.1.12 Virtual Environment Situation Awareness Rating System

General description – The Virtual Environment Situation Awareness Rating System (VESARS) is a software tool that measures SA in virtual and real-world training environments. It has three metrics: (1) real-time SA queries of individual, shared, and team SA, (2) real-time SA behavior ratings of individual and team SA by subject matter experts, and (3) real-time SA communication ratings also made by subject matter experts (Strater et al., 2013).

Strengths and limitations – The VESARS has been integrated into the Virtual Battlespace 2 training suite. Scielzo et al. (2010) used VESARS to measure the SA of two squadron leaders during a training game. They concluded that VESARS could be successfully used real time if the raters were very familiar with the rating system.

Data requirements – Well-defined queries and well-trained raters.

Thresholds – 0% to 100%.

Sources

- Scielzo, S., Davis, F., Riley, J.M., Hyatt, J., Lampton, D., and Merlo, J. Leveraging serious games and advanced training technologies for enhanced cognitive skill development. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 2408–2412, 2010.
- Strater, L., Riley, J., Davis, F., Procci, K., Bowers, C., Beidel, D., Brunnell, B., Proaps, A., Sankaranarayanan, G., Li, B., De, S., and Cao, C.G.L. Me and my VE, part 2. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 2127–2131, 2013.
-

3.2 Subjective Measures of SA

Subjective measures of SA share many of the advantages and limitations of subjective measures of workload discussed in Section 3.2. Advantages include: inexpensive, easy to administer, and high face validity. Disadvantages include: inability to measure what the subject cannot describe well in words and requirement for well-defined questions.

3.2.1 China Lake Situational Awareness

General description – The China Lake Situational Awareness (CLSA) is a five-point rating scale (see Table 3.2) based on the Bedford Workload Scale. It was designed at the Naval Air Warfare Center at China Lake to measure SA in flight (Adams, 1998).

Strengths and limitations – Jennings et al. (2004) reported a significant increase in CLSA ratings associated with a Tactile Situational Awareness System that provided tactile cues for maintaining aircraft position. The participants were 11 pilots flying a Bell 205 Helicopter.

Bruce Hunn (personal communication, 2001) argued that the CLSA “fails to follow common practice in rating scale design, does not provide diagnostic results and in general, is unsuitable for assessing SA in a test environment.” He argues that the scale cannot measure SA since it does not include the three components of SA: perception, comprehension, or projection. Further, Hunn argues that the terminology is not internally consistent, includes multiple dimensions and compound questions, and has not been validated.

Data requirements – Points in the flight during which the aircrew are asked to rate their SA using the CLSA Rating Scale must not compromise safety.

Thresholds – 1 (very good) to 5 (very poor).

TABLE 3.2

China Lake SA Rating Scale

SA Scale Value	Content
VERY GOOD 1	Full knowledge of a/c energy state/tactical environment/mission; Full ability to anticipate/accommodate trends
GOOD 2	Full knowledge of a/c energy state/tactical environment/mission; Partial ability to anticipate/accommodate trends; No task shedding
ADEQUATE 3	Full knowledge of a/c energy state/tactical environment/mission; Saturated ability to anticipate/accommodate trends; Some shedding of minor tasks
POOR 4	Fair knowledge of a/c energy state/tactical environment/mission; Saturated ability to anticipate/accommodate trends; Shedding of all minor tasks as well as many not essential to flight safety/mission effectiveness
VERY POOR 5	Minimal knowledge of a/c energy state/tactical environment/mission; Oversaturated ability to anticipate/accommodate trends; Shedding of all tasks not absolutely essential to flight safety/mission effectiveness

Sources

- Adams, S. Practical considerations for measuring situational awareness. Proceedings for the 3rd Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment, 157–164, 1998.
- Jennings, S., Craig, G., Cheung, B., Rupert, A., and Schultz, K. Flight-test of a tactile situational awareness system in a land-based deck landing task. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 142–146, 2004.

3.2.2 Crew Awareness Rating Scale

General description – The Crew Awareness Rating Scale (CARS) has eight scales (see Table 3.3) that are rated from 1 (the ideal case) to 4 (the worst case).

Strengths and limitations – McGuinness and Foy (2000) reported that the CARS had been successfully used in studies of airline pilot eye movements, automation trust on flight decks of commercial aircraft, and military command and control. Foy and McGuinness (2000) reported significant differences between pilots provided with a Traffic Collision Avoidance System (TCAS) and those without – but only for the resolution component. McGuinness et al. (2000) used CARS to compare SA between conventional and digital command and control displays. Understanding the big picture had significantly higher SA in the digital condition. McGuinness and Ebbage (2002) used CARS to evaluate the SA of seven Commanding Officer/Operations Officer teams in the Royal Military College of Science. All seven

TABLE 3.3

Definitions of CARS Rating Scales

Perception – the assimilation of new information
1. The content of perception – is it reliable and accurate?
2. The processing of perception – is it easy to maintain?
Comprehension – the understanding of information in context
3. The content of comprehension – is it reliable and accurate?
4. The processing of comprehension – is it easy to maintain?
Projection – the anticipation of possible future developments
5. The content of projection – is it reliable and accurate?
6. The processing of projection – is it easy to maintain?
Integration – the synthesis of the above with one’s course of action
7. The content of integration – is it reliable and accurate?
8. The processing of integration – is it easy to maintain?

teams completed two two-hour land reconnaissance missions; one with a standard radio and one with a digital map and electronic text messaging. Content ratings gradually improved over the exercise. The processing ratings were not significantly different between the standard and digitized versions of the mission. Content, however, was higher in the digital version except for the Commanding Officers who rated comprehension higher using the standard radio.

CARS has also been used to evaluate the relationships among workload, teamwork, SA, and performance (Berggren et al. 2011). They reported that seven of the eight CARS questions were identified in a Principal Component Analysis (PCA) as a latent factor in teamwork. The other factors were performance, workload, and teamwork. The participants were 18 two-person teams of university students fighting a virtual fire.

In a recent aviation study, Stelzer et al. (2013) reported no differences in SA as measured by CARS with and without a block occupancy display. The participants were 11 former or current air traffic controllers.

Data requirements – Use of the standard CARS rating scales.

Sources

- Berggren, P., Prytz, E., Johansson, B., and Nahlinder, S. The relationship between workload, teamwork, situation awareness, and performance in teams: A micro-world study. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 851–855, 2011.
- Foy, L., and McGuinness, B. Implications of cockpit automation for crew situational awareness. Proceedings of the 1st Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 101–106, 2000.

- McGuinness, B., and Ebbage, L. *Assessing Human Factors in Command and Control: Workload and Situational Awareness Metrics*. Bristol, United Kingdom: BAE Systems Advanced Technology Centre, May 2002.
- McGuinness, B., and Foy, L. A subjective measure of SA: The Crew Awareness Rating Scale (CARS). Proceedings of the Human Performance, Situation Awareness, and Automation Conference, 286–291, 2000.
- McGuinness, B., Foy, L., and Forsey, T. Battlespace digitization: SA issues for commanders. Proceedings of the First Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 125, 2000.
- Stelzer, E.K.M., Chong, R.S., Stevens, R.K., and Nene, V.D. Controller use of a block occupancy-based surface surveillance display for surface management. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 51–55, 2013.

3.2.3 Crew Situational Awareness

General description – Mosier and Chidester (1991) developed a method for measuring situational awareness of air transport crews. Expert observers rate crew coordination performance and identify and rate performance errors (type 1, minor errors; type 2, moderately severe errors; and type 3, major, operationally significant errors). The experts then develop information transfer matrices identifying time and source of item requests (prompts) and verbalized responses. Information is then classified into decision or nondecision information.

Strengths and limitations – The method was sensitive to type of errors and decision prompts.

Data requirements – The method requires open and frequent communication among aircrew members. It also requires a team of expert observers to develop the information transfer matrices.

Thresholds – Not stated.

Source

Mosier, K.L., and Chidester, T.R. Situation assessment and situation awareness in a team setting. In R.M. Taylor (Ed.) *Situation Awareness in Dynamic Systems (IAM Report 708)*. Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

3.2.4 Mission Awareness Rating Scale (MARS)

General description – The Mission Awareness Rating Scale (MARS) is an eight-question rating scale designed for use by infantry personnel (see Table 3.4).

Strengths and limitations – Matthews and Beal (2002) asked 16 cadets attending the U.S. Military Academy to provide MARS ratings during a battalion level exercise. Platoon leaders rated their SA higher than squad leaders.

Data requirements – Completion of the ratings for 8 questions.
Thresholds – 0 to 3 for content and workload.

Source

Matthews, M.D., and Beal, S.A. *Assessing Situation Awareness in Field Training Exercises (Research Report 1795)*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, September 2002.

3.2.5 Human Interface Rating and Evaluation System

General description – The Human Interface Rating and Evaluation System (HiRes) is a generic judgment-scaling technique developed by Budescu et al. (1986).

Strengths and limitations – HiRes has been used to evaluate SA (Fracker and Davis, 1990). These authors reported a significant effect on the number of enemy aircraft in a simulation and HiRes rating.

Data requirements – HiRes ratings are scaled to sum to 1.0 across all the conditions to be rated.

Thresholds – 0 to 1.0.

Sources

Budescu, D.V., Zwick, R., and Rapoport, A. A comparison of the Eigen value and the geometric mean procedure for ratio scaling. *Applied Psychological Measurement* 10(1): 69–78, 1986.

Fracker, M.L., and Davis, S.A. Measuring operator situation awareness and mental workload. Proceedings of the 5th Mid-Central Ergonomics/Human Factors Conference, 23–25, 1990.

3.2.6 Situation Awareness for SHAPE

General description – As part of EUROCONTROL's Solution for Human-Automation Partnerships in European ATM (SHAPE) project, the Situation Awareness for SHAPE (SASHA) was developed. There are two forms of SASHA: (1) SASHA on-Line (SASHA_L) is a set of queries and (2) SASHA Questionnaire (SASHA_Q) consists of questions related to elements of SA identified by Air Traffic Controllers (Straeter and Woldring, 2003). For SASHA_L, a subject matter expert views an Air Traffic Controller's

TABLE 3.4

Mission Awareness Rating Scales

Instructions. Please answer the following questions about the mission you just completed.

Your answers to these questions are important in helping us evaluate the effectiveness of this training exercise. Check the response that best applies to your experience.

The first four questions deal with your ability to detect and understand important cues present during the mission.

1. Please rate your ability to identify mission-critical cues in this mission.

- very easy – able to identify all cues
- fairly easy – could identify most cues
- somewhat difficult – many cues hard to identify
- very difficult – had substantial problems identifying most cues

2. How well did you understand what was going on during the mission?

- very well – fully understood the situation as it unfolded
- fairly well – understood most aspects of the situation
- somewhat poorly – had difficulty understanding much of the situation
- very poorly – the situation did not make sense to me

3. How well could you predict what was about to occur next in the mission?

- very well – could predict with accuracy what was about to occur
- fairly well – could make accurate predictions most of the time
- somewhat poor – misunderstood the situation much of the time
- very poor – unable to predict what was about to occur

4. How aware were you of how to best achieve your goals during this mission?

- very aware – knew how to achieve goals at all times
- fairly aware – knew most of the time how to achieve mission goals
- somewhat unaware – was not aware of how to achieve some goals
- very unaware – generally unaware of how to achieve goals

The last four questions ask how difficult it was for you to detect and understand important cues present during the mission.

5. How difficult – in terms of mental effort required – was it for you to identify or detect mission-critical cues in the mission?

- very easy – could identify relevant cues with little effort
- fairly easy – could identify relevant cues, but some effort required
- somewhat difficult – some effort was required to identify most cues
- very difficult – substantial effort required to identify relevant cues

6. How difficult – in terms of mental effort – was it to understand what was going on during the mission?

- very easy – understood what was going on with little effort
- fairly easy – understood events with only moderate effort
- somewhat difficult – hard to comprehend some aspects of situation
- very difficult – hard to understand most or all aspects of situation

7. How difficult – in terms of mental effort – was it to predict what was about to happen during the mission?

- very easy – little or no effort needed
- fairly easy – moderate effort required

(Continued)

TABLE 3.4 (CONTINUED)

Mission Awareness Rating Scales

___ somewhat difficult – many projections required substantial effort

___ very difficult – substantial effort required on most or all projections

8. How difficult – in terms of mental effort – was it to decide on how to best achieve mission goals during this mission?

___ very easy – little or no effort needed

___ fairly easy – moderate effort required

___ somewhat difficult – substantial effort needed on some decisions

___ very difficult – most or all decisions required substantial effort (Matthews and Beal, 2002, p. A-1–A-2).

displays and in real time asks SA-related questions. The questions are rated by the subject matter expert on their *operational* importance (Straeter and Woldring, 2003). The time for an Air Traffic Controller to respond to the question is rated as OK, too long, or too short. An example of a question is “Which aircraft needs to be transferred next?” (Straeter and Woldring, 2003, p. 50). SASHA_Q is a self-rating questionnaire. An example is presented in Figure 3.3.

Strengths and limitations – Dehn (2008) described the steps taken in the development of the SASHA: (1) literature review to obtain initial set of items, (2) requirement-based review for easy administration, ease of understanding, consistent format, and scoring key provided, (3) collection of expert feedback, and (4) initial empirical study. The study participants were 24 active air traffic controllers who completed the SASHA. Items were eliminated from the questionnaire if they reduced internal consistency or were redundant. Jipp and Papenfuss (2011) reported problems replicating the scales as well as finding results contrary to previous research. Their participants were 12 professional tower controllers working in teams during an airspace simulation. The SHAPE Teamwork Questionnaire (STQ) was completed after each of the three simulation runs.

In the previous working period(s) ...

	never	seldom	sometimes	often	more often	very often	always
1) ... I was ahead of the traffic.	0	1	2	3	4	5	6
2) ... I started to focus on a single problem or a specific area of the sector.	0	1	2	3	4	5	6

FIGURE 3.3
SASHA (Dehn, 2008, p. 138).

Data requirements – Data are responses to questionnaire items. A scoring key for SASHA is available from EUROCONTROL Headquarters. It includes scores for both the inverted and noninverted questions. There is also a user's guide (EUROCONTROL Edition Number 0.1, 30 July 2007).

Thresholds – The data are on an ordinal scale and must be treated accordingly when statistical analysis is applied to the data. Non-parametric statistics may be the most appropriate analysis method.

Sources

Dehn, D.M. Assessing the impact of automation on the air traffic controller: The SHAPE questionnaires. *Air Traffic Control Quarterly* 16(2): 127–146, 2008.

EUROCONTROL. *The New SHAPE Questionnaires: A User's Guide*. Edition Number 0.1. 30 July 2007. http://www.eurocontrol.int/humanfactors/public/standard_page/15_newsletter_SHAPE.html

Jipp, M., and Papenfuss, A. Reliability and validity of the SHAPE Teamwork Questionnaire. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 115–119, 2011.

Straeter, O., and Woldring, M. *The Development of Situation Awareness Measures in ATM Systems (HRS/HSP-005-REP-01)*. Brussels: EATMP Infocenter. EUROCONTROL Headquarters, June 27, 2003.

3.2.7 Situation Awareness Behavioral Rating Scale (SABARS)

General description – The Situation Awareness Behavioral Rating Scale (SABARS) is a 28-question rating scale measuring SA of infantry personnel (see Table 3.5). It is completed by observers.

Strengths and limitations – Matthews and Beal (2002) asked six infantry officers and four infantry noncommissioned officers to rate 16 U.S. Military Academy during a battalion level exercise. There were no significant differences in the ratings given by the commissioned and noncommissioned officers. In an earlier study, Matthews et al. (2000) recommended use of three SA measures concurrently: SABARS, SAGAT, and a Participant Situation Awareness Questionnaire (PSAQ).

Data requirements – Completion of the ratings for 28 questions.

Thresholds – 1 to 6 for each question.

TABLE 3.5
 Situation Awareness Behavioral Rating Scale

Rating Scale:

1. Very Poor
2. Poor
3. Borderline
4. Good
5. Very good
6. Not applicable/can't say

RATING (Circle one)

1. Sets appropriate levels of alert	1 2 3 4 5 6
2. Solicits information from subordinates	1 2 3 4 5 6
3. Solicits information from civilians	1 2 3 4 5 6
4. Solicits information from commanders	1 2 3 4 5 6
5. Effects coordination with other platoon/squad leaders	1 2 3 4 5 6
6. Communicates key information to commander	1 2 3 4 5 6
7. Communicates key information to subordinates	1 2 3 4 5 6
8. Communicates key information to other platoon/squad leaders	1 2 3 4 5 6
9. Monitors company net	1 2 3 4 5 6
10. Assesses information received	1 2 3 4 5 6
11. Asks for pertinent intelligence information	1 2 3 4 5 6
12. Employs squads/fire teams tactically to gather needed information	1 2 3 4 5 6
13. Employs graphic or other control measures for squad execution	1 2 3 4 5 6
14. Communicates to squads/fire teams, situation and commander's intent	1 2 3 4 5 6
15. Utilizes a standard reporting procedure	1 2 3 4 5 6
16. Identifies critical mission tasks to squad/fire team leaders	1 2 3 4 5 6
17. Ensures avenues of approach are covered	1 2 3 4 5 6
18. Locates self at vantage point to observe main effort	1 2 3 4 5 6
19. Deploys troops to maintain platoon/squad communications	1 2 3 4 5 6
20. Uses assets to effectively assess environment	1 2 3 4 5 6
21. Performs a leader's recon to assess terrain and situation	1 2 3 4 5 6
22. Identifies observation points, avenues of approach, key terrain, obstacles, cover and concealment	1 2 3 4 5 6
23. Assesses key finds and unusual events	1 2 3 4 5 6
24. Discerns key/critical information from maps, records, and supporting site information	1 2 3 4 5 6
25. Discerns key/critical information from reports received	1 2 3 4 5 6
26. Projects future possibilities and creates contingency plans	1 2 3 4 5 6
27. Gathers follow up information when needed	1 2 3 4 5 6
28. Overall Situation Awareness Rating	1 2 3 4 5 6

Matthews and Beal (2002, p. B-1).

Sources

Matthews, M.D., and Beal, S.A. *Assessing Situation Awareness in Field Training Exercises (Research Report 1795)*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, September 2002.

Matthews, M.D., Pleban, R.J., Endsley, M.R., and Strater, L.D. Measures of infantry situation awareness for a virtual MOUT environment. *Proceedings of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference*, 262–267, 2000.

3.2.8 Situation Awareness Control Room Inventory

General description – The Situation Awareness Control Room Inventory (SACRI) was developed to assess SA of operators in nuclear power plants. It has four sets of questions. The first set compares the current situation with that of the recent past (19 questions for the primary circuit and 20 questions for the secondary circuit). The second set of questions compares the current situation with normal operations but uses the same questions as the first set. The third set requires the operator to predict the future state of the system over the next few minutes. Again, the same questions are used but are rated as: “(1) increase/same, (2) decrease/same, (3) increase/same/decrease, and (4) increase in more than one/increase in one/same/decrease in one/decrease in more than one/drift in both directions” (Hogg et al., 1995, p. 2413).

Strengths and limitations – Hogg et al. (1995) reported the results of four nuclear power plant simulator studies and concluded that SACRI provides a sensitive measure of SA but requires experienced operators.

Data requirements – Participant responses are compared to actual plant state.

Thresholds – Not stated.

Source

Hogg, D.N., Folles, K., Strand-Volden, F., and Torralba, B. Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics* 38(11): 2394–2413, 1995.

3.2.9 Situational Awareness Rating Technique (SART)

General description – The Situational Awareness Rating Technique (SART) (see Figure 3.4) (Taylor, 1990) is a questionnaire method that concentrates on measuring the operator’s knowledge in three areas: (1) demands on attentional

resources, (2) supply of attentional resources, and (3) understanding of the situation (see Figures 3.3 and 3.4 and Table 3.6). The reason that SART measures three different components (there is also a 10-dimensional version) is that the SART developers proposed that, like workload, SA is a complex construct; therefore, to measure SA in all its aspects, separate measurement dimensions are required. Because information processing and decision making are inextricably bound with SA (since SA involves primarily cognitive rather than physical workload), SART has been tested in the context of Rasmussen’s Model of skill-, rule-, and knowledge-based behavior. Selcon and Taylor (1989) conducted separate studies looking at the relationship between SART and rule- and knowledge-based decisions, respectively. The results showed that SART ratings appear to provide diagnosticity in that they were significantly related to performance measures of the two types of decision making.

Strengths and limitations – SART is a subjective measure and, as such, suffers from the inherent reliability problems of all subjective measures. Taylor and Selcon (1991) state: “There remains considerable scope for scales development, through description improvement, interval justification and the use of conjoint scaling techniques to condense multi-dimensional ratings into

		LOW							HIGH								
		1	2	3	4	5	6	7	1	2	3	4	5	6	7		
Demand	Instability of Situation																
	Variability of Situation																
	Complexity of Situation																
Supply	Arousal																
	Spare Mental Capacity																
	Concentration																
	Division of Attention																
Under	Information Quantity																
	Information Quality																
	Familiarity																

FIGURE 3.4
SART scale.

TABLE 3.6
Definitions of SART Rating Scales

Demand on Attentional Resources
Instability: Likelihood of situation changing suddenly
Complexity: Degree of complication of situation
Variability: Number of variables changing in situation
Supply of Attentional Resources
Arousal: Degree of readiness for activity
Concentration: Degree of readiness for activity
Division: Amount of attention in situation
Space capacity: Amount of attention left to spare for new variables
Understanding of the Situation
Information quantity: Amount of information received and understood
Information quality: Degree of goodness of information gained

From Taylor and Selcon (1991, p. 10).

a single SA score” (p. 11). These authors further state that “The diagnostic utility of the Attentional Supply constructs has yet to be convincingly demonstrated” (p. 12).

The strengths are that SART is easily administered and was developed in three logical phases: (1) scenario generation, (2) construct elicitation, and (3) construct structure validation (Taylor, 1989). SART has been prescribed for comparative system design evaluation (Taylor and Selcon, 1991). It has been used in aviation, cybersecurity, military operations, submarines, and surface transportation. It has also been adapted to measure cognitive compatibility.

Aviation. SART is sensitive to differences in performance of aircraft attitude recovery tasks and learning comprehension tasks (Selcon and Taylor, 1991; Taylor and Selcon, 1990). SART is also sensitive to pilot experience (Selcon et al., 1991), timeliness of weather information in a simulated flight task (Bustamante et al., 2005), and field of view and size of a flight display in a fixed-base flight simulator (Stark et al., 2001). Wilson et al. (2002) reported a significant difference in SART scores associated with alternate Head Up Display symbologies. Their participants were 27 pilots performing taxi maneuvers in a fixed-based simulator. Boehm-Davis et al. (2010) reported significantly less awareness for pilots in using Data Comm than using voice. The participants were 24 transport pilots. The data were collected in a simulator. Burke et al. (2016) used SART to evaluate the SA of 12 pilots using a trajectory optimization system in a Piaggio Avanti P180.

Selcon et al. (1992) used SART to evaluate the effectiveness of visual, auditory, or combined cockpit warnings. Demand was significantly greater for the visual than for the auditory or the combined cockpit warnings. Neither supply nor understanding ratings were significantly different, however, across these conditions. Similarly, Selcon et al. (1996) reported significantly higher

SART scores when a launch success zone display was available to pilots during a combat aircraft simulation. Understanding, information quantity, and information quality were also significantly higher with this display. There were no effects on Demand or Supply ratings. Strybel et al. (2007) reported that nine instrument-rated pilots responded to seven probe questions after a simulated ILS approach. There were significant predictions from responses to three of these questions and SART combined score, SART understanding, and SART demand.

Vidulich et al. (1995) reported increased difficulty of a PC-based flight simulation increased the Demand Scale on the SART but not the Supply or Understanding Scales. Providing additional information increased the Understanding Scale but not the Demand or Supply Scales. Crabtree et al. (1993) reported that the overall SART rating discriminated SA in a simulated air-to-ground attack while a simple overall SA rating did not. The test-retest reliability of the overall SART was not significant. See and Vidulich (1997) reported significant effects of target and display type on SART. The combined SART as well as the supply and understanding scales were significantly correlated to workload ($r = -0.73, -0.75, \text{ and } -0.82$, respectively).

In a large experiment (80 men and 180 women), Fink and Major (2000) compared SART and SALIENT related to a helicopter flight simulation game (Werewolf v Comanche). The authors reported that the SALIENT had the better measurement characteristics but SART enabled them to assess the interaction between the game and the operator. In another helicopter (Blackhawk Helicopter simulator) Casto and Casali (2010) reported a significant effect of workload manipulation (decreasing visibility, increasing number of maneuvers, and number of communications) and communication signal quality on SART ratings. Brown and Galster (2004) reported no significant effect of varying the reliability of automation in a simulated flight task on SART. Their participants were eight male pilots.

Verma et al. (2010) used SART to evaluate temporal separation between commercial aircraft as well as straight in and slewed approaches. Their participants were three retired commercial pilots. Due to the small amount of data, no statistical analyses were performed. However, the authors concluded that there were no differences between 5 and 10 second separation or type of path.

In an Air Traffic Control scenario, Vu et al. (2009) reported significantly higher SA in a low air traffic density scenario than in a high traffic density scenario. Taylor et al. (1995) reported significant differences in both 3-D and 10-D SART ratings in an ATC task. Only three scales of the 10-D SART did not show significant effects as the number of aircraft being controlled changed (Information Quantity, Information Quality, and Familiarity). Durso et al. (1999) reported that Air Traffic Controllers made little distinction between understanding and SA.

Gregg et al. (2012) evaluated conflict detection automation among 13 recently retired Air Traffic Controllers using an ATC simulator. Participants were divided into two groups based on their rating of the complexity of four automation systems. Group 1 had significantly higher SART ratings for every study condition than Group 2.

Cyber Security. In a cyber threat detection task, Giacobe (2013) found no significant effects of experience on SART ratings but there were significant differences between text and graphic presentation of information on three SART scales: Attention Division, Information Quantity, and Familiarity.

Military Operations. Schuster et al. (2012) reported a significant interaction in SART scores such that a high information first group had higher SART scores on the third mission than on the second mission. The task was to identify the affiliation of soldiers in a simulated Military Operations in Urban Terrain (MOUT). The task was done in human-robot teams.

Submarines. Huf and French (2004) reported SART indicating low understanding of the situation in a Virtual Submarine Environment while SAGAT indicated high understanding as evidenced by correct responses. Loft et al. (2015) compared SPAM, SAGAT, ATWIT, NASA TLX, and SART ratings from 117 undergraduates performing three submarine tasks: contact classification, closest point of approach, and emergency surface. SPAM did not significantly correlate with SART but did with ATWIT and NASA TLX.

Surface Transportation. Read and Sallam (2017) used SART to compare SA in a Virtual Reality Head-Mounted Display, real world driving, and a flat screen to train drivers. They reported no significant differences among the three training environments.

Cognitive Compatibility. SART was modified to measure cognitive compatibility (CC-SART) to assess the effects of color coding in the Gripen fighter aircraft (Derefeldt et al., 1999). CC-SART has three primary and 10 subsidiary scales. Only the three primary scales were used: (1) level of processing, (2) ease of reasoning, and (3) activation of knowledge. Participants were seven Swedish fighter pilots performing a simulated tracking of an adversarial aircraft using the Head Up Display and detecting a target on the head down display. The CC-SART Index was lowest for the monochromatic displays and highest for the dichrome color displays. However, the shortest RT to detecting the head down target occurred with polychromatic displays. Parasuraman et al. (2009) used CC-SART to evaluate three automation modes (manual, static automation, automated target recognition) on supervision of multiple uninhabited vehicles. These authors reported that the score for level of processing was significantly lower than ease of reasoning.

Data requirements – Data are on an ordinal scale; interval or ratio properties cannot be implied.

Thresholds – Not stated.

Sources

- Boehm-Davis, D.A., Gee, S.K., Baker, K., and Medina-Mora, M. Effect of party line loss and delivery format on crew performance and workload. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 126–130, 2010.
- Brown, R.D., and Galster, S.M. Effects of reliable and unreliable automation on subjective measures of mental workload, situation awareness, trust and confidence in a dynamics flight task. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 147–151, 2004.
- Burke, K.A., Wing, D.J., and Haynes, M. Flight test assessments of pilot workload, system usability, and situation awareness of TASAR. Proceedings of the Human Factors and Ergonomics Society 60th Annual Meeting, 61–65, 2016.
- Bustamante, E.A., Fallon, C.K., Bliss, J.P., Bailey, W.R., and Anderson, B.L. Pilots' workload, situation awareness, and trust during weather events as a function of time pressure, role assignment, pilots' rank, weather display, and weather system. *International Journal of Applied Aviation Studies* 5(2): 348–368, 2005.
- Casto, L.K.L., and Casali, J.G. Effect of communications headset, hearing ability, flight workload, and communications signal quality on pilot performance in an Army Black Hawk helicopter simulator. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, 80–84, 2010.
- Crabtree, M.S., Marcelo, R.A.Q., McCoy, A.L., and Vidulich, M.A. An examination of a subjective situational awareness measure during training on a tactical operations trainer. Proceedings of the 7th International Symposium on Aviation Psychology, 891–895, 1993.
- Derefeldt, G., Skinnars, Ö., Alfredson, J., Eriksson, L., Andersson, P., Westlund, J., Berggrund, U., Holmberg, J., and Santesson, R. Improvement of tactical situation awareness with color-coded horizontal-situation displays in combat aircraft. *Displays* 20(4): 171–184, 1999.
- Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D., and Manning, C.A. *Situation Awareness as a Predictor of Performance in En Route Air Traffic Controllers (DOT/FAA/AM-99/3)*. Washington, DC: Office of Aviation Medicine, January 1999.
- Fink, A.A., and Major, D.A. Measuring situation awareness: A comparison of three techniques. Proceedings of the 1st Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, 256–261, 2000.
- Giacobe, N.A. A picture is worth a thousand words. Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, 172–176, 2013.
- Gregg, S., Martin, L., Homola, J., Lee, P., Mercer, J., Brasil, C., Cabrall, C., and Lee, H. Shifts in air traffic controllers' situation awareness during high altitude mixed equipage operations. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 95–99, 2012.
- Huf, S., and French, H.T. Situation awareness in a networked virtual submarine. Proceedings of the Human Factors and Ergonomics 48th Annual Meeting, 663–667, 2004.

- Loft, S., Bowden, V., Braithwaite, J., Morrell, D.B., Huf, S., and Durso, F.T. Situation awareness measures for simulated submarine track management. *Human Factors* 57(2): 298–310, 2015.
- Parasuraman, R., Cosenzo, K.A., and Visser, E.D. Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, 21(2): 270–297, 2009.
- Read, J.M., and Sallam, J.J. Task performance and situation awareness with a virtual reality head-mounted display. Proceedings of the Human Factors and Ergonomics 61st Annual Meeting, 2105–2109, 2017.
- Schuster, D., Keebler, J.R., Jentsch, F., and Zuniga, J. Comparison of SA measurement techniques in a human-robot team task. Proceedings of the Human Factors Society 56th Annual Meeting, 1713–1717, 2012.
- See, J.E., and Vidulich, M.A. Assessment of computer modeling of operator mental workload during target acquisition. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, 1303–1307, 1997.
- Selcon, S.J., Hardiman, T.D., Croft, D.G., and Endsley, M.R. A test-battery approach to cognitive engineering: To meta-measure or not to meta-measure, that is the question!. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, 228–232, 1996.
- Selcon, S.J., and Taylor, R.M. Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. AGARD Conference Proceedings No. 478, Neuilly-sur-Seine, France, 1989.
- Selcon, S.J., and Taylor, R.M. Decision support and situational awareness. In R.M. Taylor (Ed.) *Situational Awareness in Dynamic Systems (IAM Report 708)*. Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.
- Selcon, S.J., Taylor, R.M., and Koritsas, E. Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation. Proceedings of the Human Factors Society 35th Annual Meeting, 62–66, 1991.
- Selcon, S.J., Taylor, R.M., and Shadrake, R.A. Multi-modal: Pictures, words, or both? Proceedings of the Human Factors Society 36th Annual Meeting, 57–61, 1992.
- Stark, J.M., Comstock, J.R., Prinzel, L.J., Burdette, D.W., and Scerbo, M.W. A preliminary examination of situation awareness and pilot performance in a synthetic vision environment. Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, 40–43, 2001.
- Strybel, T.Z., Vu, K.L., Dwyer, J.P., Kraft, J., Ngo, T.K., Chambers, V., and Garcia, F.P. Predicting perceived situation awareness of low altitude aircraft in terminal airspace using probe questions. *Human-Computer Interaction. Interaction Design and Usability. Lecture Notes in Computer Science* (volume 4550). Berlin: Springer, July 2007.
- Taylor, R.M. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. Proceedings of the NATO Advisory Group for Aerospace Research and Development (AGARD) Situational Awareness in Aerospace Operations Symposium (AGARD-CP-478), October 1989.
- Taylor, R.M. *Situational Awareness: Aircrew Constructs for Subject Estimation (IAM-R-670)*. Farnborough, UK: Institute of Aviation Medicine, 1990.
- Taylor, R.M., and Selcon, S.J. Understanding situational awareness. Proceedings of the Ergonomics Society's 1990 Annual Conference, 105–111, 1990.

- Taylor, R.M., and Selcon, S.J. Subjective measurement of situational awareness. In R.M. Taylor (Ed.) *Situational Awareness in Dynamic Systems (IAM Report 708)*. Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.
- Taylor, R.M., Selcon, S.J., and Swinden, A.D. Measurement of situational awareness and performance: A unitary SART index predicts performance on a simulated ATC task. Proceedings of the 21st Conference of the European Association for Aviation Psychology. Chapter 41, 1995.
- Verma, S., Lozito, S.C., Ballinger, D.S., Trot, G., Hardy, G.H., Panda, R.C., Lehmer, R.D., and Kozon, T.E. *Preliminary Human-in-the-Loop Assessment of Procedures for Very-Closely-Spaced Parallel Runways (NASA/TM-2010-216026)*. Moffett Field, California: NASA Ames Research Center, April 2010.
- Vidulich, M.A., McCoy, A.L., and Crabtree, M.S. The effect of a situation display on memory probe and subjective situational awareness metrics. Proceedings of the 8th International Symposium on Aviation Psychology, 765–768, 1995.
- Vu, K.L., Minakata, K., Nguyen, J., Kraut, J., Raza, H., Battiste, V., and Strybel, T.Z. Situation awareness and performance of student versus experienced air traffic controllers. In M.J. Smith and G. Salvendy (Eds.) *Human Interface, Part II, HCII 2009*, LNCS5618 (pp. 865–874), 2009.
- Wilson, J.R., Hooey, B.L., Foyle, D.C., and Williams, J.L. Comparing pilots' taxi performance, situation awareness and workload using command-guidance, situation-guidance and hybrid head-up display symbologies. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, 16–20, 2002.

3.2.10 Situational Awareness Subjective Workload Dominance

General descriptions – The Situation Awareness Subjective Workload Dominance Technique (SA-SWORD) uses judgment matrices to assess SA.

Strengths and limitations – Fracker and Davis (1991) evaluated alternate measures of SA on three tasks: (1) flash detection, (2) color identification, and (3) location. Ratings were made of awareness of object location, color, flash, and mental workload. All ratings were collected using a paired-comparisons technique. Color inconsistency decreased SA and increased workload. Flash probability had no significant effects on the ratings. Ruff et al. (2000) reported significantly higher SA for management by consent automation mode than for manual control or management by exception. The task was the control of one, two, or four remotely operated vehicles. However, Snow and Reising (2000) found no significant correlation between SAGAT and SA-SWORD and further only SA-SWORD showed statistically significant effects of visibility and synthetic terrain type in a simulated flight.

Data requirements – There are three required steps: (1) a rating scale listing all possible pairwise comparisons of the tasks performed must be completed, (2) a judgment matrix comparing each task to every other task must be filled in with each participant's evaluation of the tasks, and (3) ratings must be calculated using a geometric means approach.

Thresholds – Not stated.

Sources

Fracker, M.L., and Davis, S.A. *Explicit, Implicit, and Subjective Rating Measures of Situation Awareness in a Monitoring Task (AL-TR-1991-0091)*. OH: Wright-Patterson Air Force Base, October 1991.

Ruff, H.A., Draper, M.H., and Narayanan, S. The effect of automation level and decision aid fidelity on the control of multiple remotely operated vehicles. *Proceedings of the 1st Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium*, 2000.

Snow, M.P., and Reising, J.M. Comparison of two situation awareness metrics: SAGAT and SA-SWORD. *Proceedings of the IEA 2000/HFES 2000 Congress*, vol. 3, 49–52, 2000.

3.2.11 Situational Awareness Supervisory Rating Form

General descriptions – Carretta et al. (1996) developed the Situational Awareness Supervisory Rating Form to measure the SA capabilities of F-15 pilots. The form has 31 items that range from general traits to tactical employment (Table 3.7).

Strengths and limitations – Carretta et al. (1996) reported that 92.5% of the variance in peer and supervisory ratings were due to one principal component. The best predictor of the form rating was flying experience ($r = +0.704$).

TABLE 3.7

Situational Awareness Supervisory Rating Form

Rater ID#: _____ Pilot ID#: _____						
Item Ratings	Relative Ability Compared with Other F-15C Pilots					
	Acceptable		Good		Outstanding	
	1	2	3	4	5	6
General Traits						
1. Discipline						
2. Decisiveness						
3. Tactical knowledge						
4. Time-sharing ability						
5. Reasoning ability						
6. Spatial ability						
7. Flight management						
Tactical Game Plan						
8. Developing plan						
9. Executing plan						
10. Adjusting plan on the fly						

(Continued)

TABLE 3.7 (CONTINUED)

Situational Awareness Supervisory Rating Form

System Operation 11. Radar 12. TEWS 13. Overall weapons system proficiency						
Communication 14. Quality (brevity, accuracy, timeliness, completeness) 15. Ability to effectively use comm/information						
Information Interpretation 16. Interpreting VSD 17. Interpreting RWR 18. Ability to effectively use AWACS/GCI 19. Integrating overall information (cockpit displays, wingman comm, controller comm) 20. Radar sorting 21. Analyzing engagement geometry 22. Threat prioritization						
Tactical Employment-BVR Weapons 23. Targeting decision 24. Fire-point selection						
Tactical Employment-Visual Maneuvering 25. Maintain track of bogeys/friendlies 26. Threat evaluation 27. Weapons employment						
Tactical Employment-General 28. Assessing offensiveness/defensiveness 29. Lookout (VSD interpretation, RWR monitoring, visual lookout) 30. Defensive reaction (chaff, flares, maneuvering, etc.) 31. Mutual support						
Overall situational awareness ^a						
Overall fighter ability						

^a Items 1 through 31 are used for supervisory ratings. The overall fighter ability and situational awareness items are completed by both supervisors and peers. (Carretta et al., 1996, pp. 40–41).

Waag and Houck (1996) evaluated the Situational Awareness Supervisory Rating Form as one of a set of three SA rating scales. The other two were peer and self-ratings. The data were collected from 239 F-15C pilots. Reliabilities on the three scales ranged from +0.97 to +0.99. Inter-rater reliability was +0.84. Correlations between supervisory and peer ratings ranged from +0.85 to +0.87. Correlations with the self-report were smaller (+0.50 to +0.58). In an earlier publication these authors referred to the scales as the Situational Awareness Rating Scales (SARS) (Waag and Houck, 1994).

Data requirements – Supervisors and peers must make the rating.

Sources

- Carretta, T.R., Perry, D.C., and Ree, M.J. Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology* 6(1): 21–41, 1996.
- Waag, W.L., and Houck, M.R. Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space, and Environmental Medicine* 65(5, Suppl.): A13–A19, 1994.
- Waag, W.L., and Houck, M.R. Development of criterion measures of situation awareness for use in operational fighter squadrons. Proceedings of the Advisory Group for Aerospace Research and Development Conference on Situation Awareness Limitation and Enhancement in the Aviation Environment (AGARD-CP-575). AGARD, Neuilly-sur-Seine, France, 8-1-8-8, January 1996.
-

3.3 Simulation

Shively et al. (1997) developed a computational model of SA. The model has three components: (1) situational elements, i.e., parts of the environment that define the situation, (2) context-sensitive nodes, i.e., semantically-related collections of situational elements, and (3) a regulatory mechanism that assesses the situational elements for all nodes.

See and Vidulich (1997) reported that a Micro Saint model of operator SA during a simulated air-to-ground mission matched SART predictions with the closest correlation with the understanding scale of the SART.

Sources

- See, J.E., and Vidulich, M.A. Assessment of computer modeling of operator mental workload during target acquisition. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, 1303–1307, 1997.
- Shively, R.J., Brickner, M., and Silbiger, J. A computational model of situational awareness instantiated in MIDAS (Man-machine Integration Design and Analysis). Proceedings of the 9th International Symposium on Aviation Psychology, 1454–1459, 1997.