INSTITUTE FOR DEFENSE ANALYSES

DATAWorks 2023: Situation Awareness Should be Measured Scientifically: Progress Towards Evaluation Practices for Operational Testing

John T. Haman, Project Leader

Miriam E. Armstrong
Elizabeth A. Green
Janna R. Mantua
Brian D. Vickers

April 2023

IDA Document NS D-33403

Log: H 2023-000053

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-19-D-0001, Task 2299(90) "Test Science " for the Office of the Director, Operational Test and Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
The IDA Technical Review Committee was chaired by Dr. V. Bram Lillard and consisted of Dr. Kelly M. Avery, Dr. Keyla Pagan-Rivera, Dr. Brian T. Conway, and Dr. Stacey L. Allison from the Operational Evaluation Division.

For more information:
Dr. John T. Haman, Project Leader
jhaman@ida.org • (703) 845-2132

Dr. V. Bram Lillard, Director, Operational Evaluation Division
vlillard@ida.org • (703) 845-2230

Rigorous Analysis | Trusted Expertise | Service to the Nation

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-33368-NS

DATAWorks 2023: Situation Awareness Should be Measured Scientifically: Progress Towards Evaluation Practices for Operational Testing

John T. Haman, Project Leader

Miriam E. Armstrong
Elizabeth A. Green
Janna R. Mantua
Brian D. Vickers

Executive Summary

Situation Awareness (SA) plays a key role in decision making and human performance; higher operator SA is associated with increased operator performance and decreased operator errors. In the most general terms, SA can be thought of as an individual's "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future."[1]

While "situational awareness" is a common suitability parameter for systems under test, there is no standardized method or metric for quantifying SA in operational testing (OT). This leads to varied and suboptimal treatments of SA across programs and test events. Current measures of SA are exclusively subjective and paint an inadequate picture.

Future advances in system connectedness and mission complexity will exacerbate the problem. We believe that technological improvements will necessitate increases in the complexity of the warfighters' mission, including changes to team structures (e.g., integrating human teams with human-machine teams), command and control (C2) processes (e.g., expanding C2 frameworks toward joint all-domain C2), and battlespaces (e.g., overcoming integration challenges for multi-domain operations). Operational complexity increases the information needed for warfighters to maintain high SA, and assessing SA will become increasingly important and difficult to accomplish.

IDA's test science team has proposed a piecewise approach to improve the measurement of situation awareness in operational evaluations. The aim of this presentation is to promote a scientific understanding of what SA is (and is not) and encourage discussion amongst practitioners tackling this challenging problem. We will briefly introduce Endsley's model of SA, review the tradeoffs involved in some existing measures of SA, and discuss a selection of potential improvements for SA measures during OT.

---

[1] Endsley, Mica R. "Design and Evaluation for Situation Awareness Enhancement." *Proceedings of the Human Factors Society Annual Meeting* 32 (1988), 97-101.

Situation Awareness Should be Measured Scientifically: Progress Toward Evaluation Practices for Operational Testing

Miriam Armstrong, Elizabeth Green, Janna Mantua, and Brian Vickers

April 27, 2023

Institute for Defense Analyses
730 East Glebe Road ● Alexandria, Virginia 22305

A Note on Terminology…

- *Situation* Awareness vs. *Situational* Awareness

Academic Term

Colloquial Term

Briefing Overview

- Need for SA measures
- What is SA?: Endsley's three-level situation awareness model
- Existing measures of situation awareness
- How can we promote better situation awareness measurement practices in OT?

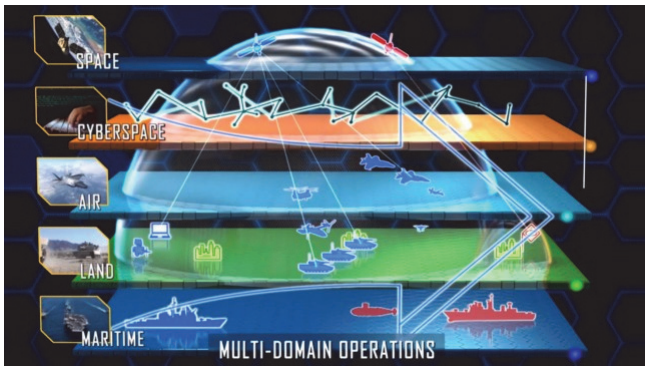
OT – Operational Testing; SA – Situation Awarenss

Situation Awareness is One of Many Human-System Integration Concepts Relevant to Operational Testing (OT)

Concept	Measures
Training	OATS
System **usability**	SUS, UMUX
Operator **workload**	AREWS, NASA-TLX
Operator **trust** of system	TOAST
Operator **situation awareness**	*[no standard within OT]*

ARWES – Air Force Flight Test Center Revised Workload Estimate Scale; HSI – Human-System Interaction; NASA-TLX – National Aeronautics and Space Administration Talk Load Index; OATS – Operational Assessment of Training Scale; OT – Operational Testing; SUS – System Usability Scale; TOAST – Trust of Automated Systems Test; UMUX – Usability Metric for User Experience

Increasing SA Challenges



Human-machine teaming and human-AI teaming

- Awareness of non-human teammates

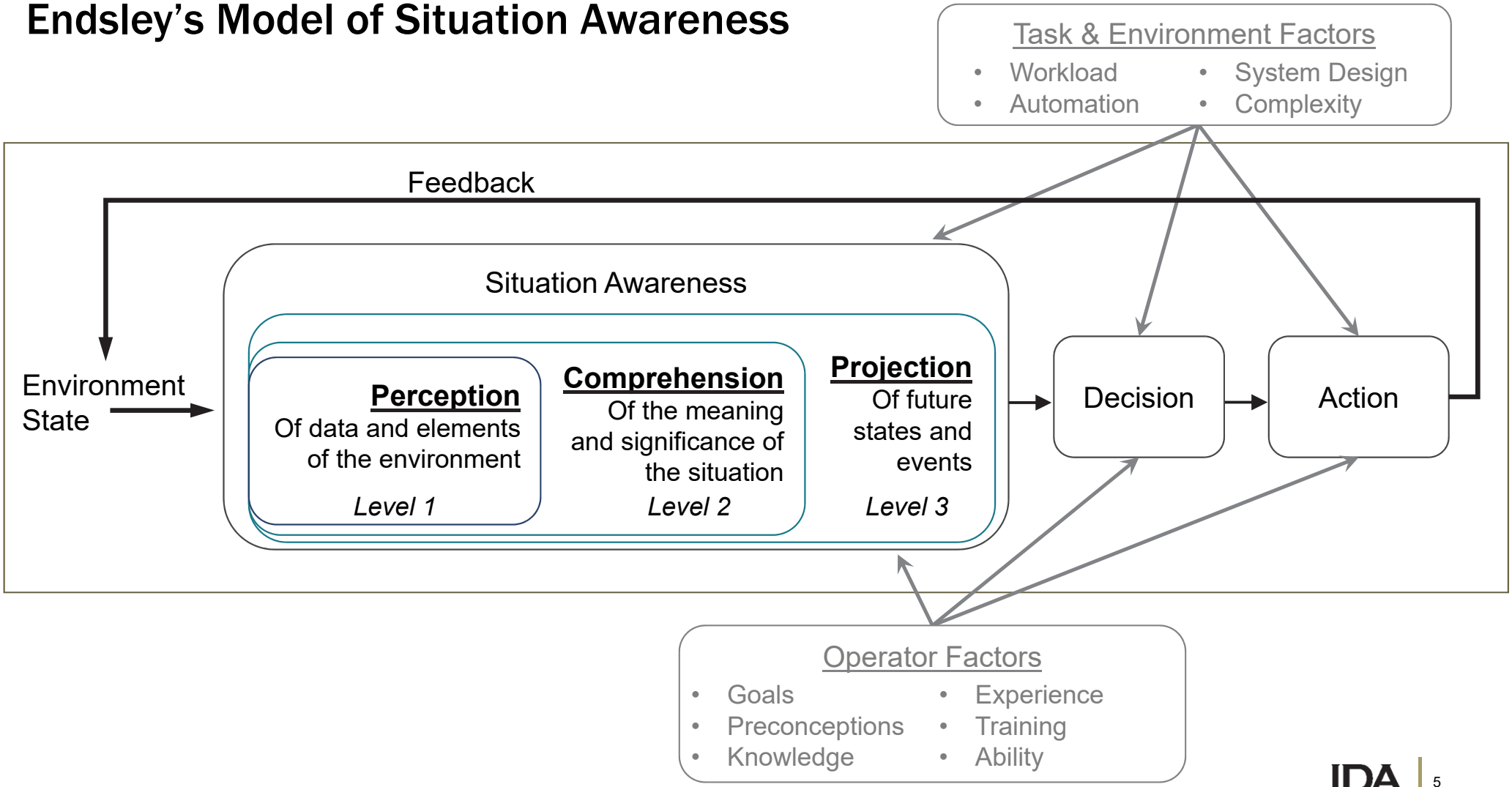
Multi-domain operations

- System effectiveness depends on SA

We need to begin training and methodological transition *now* to prepare for future programs (and problems).

AI – Artificial Intelligence; SA – Situation Awareness

Endsley's Model of Situation Awareness



Endsley's Model of Situation Awareness: Examples

Level 1 - Perception

- What terrain features are present?
- What is the current heading of the aircraft?

Level 2 - Comprehension

- Is the current route accessible for this unit?
- Is the aircraft on the intended course?

Level 3 - Projection

- How will terrain affect enemy course of action?
- Will the aircraft trajectory overlap with that of another aircraft?

How is Situation Awareness Measured?

Current OT SA measures

Custom questions

- How was your SA during this mission?
- How did the system contribute to SA?

Pros:

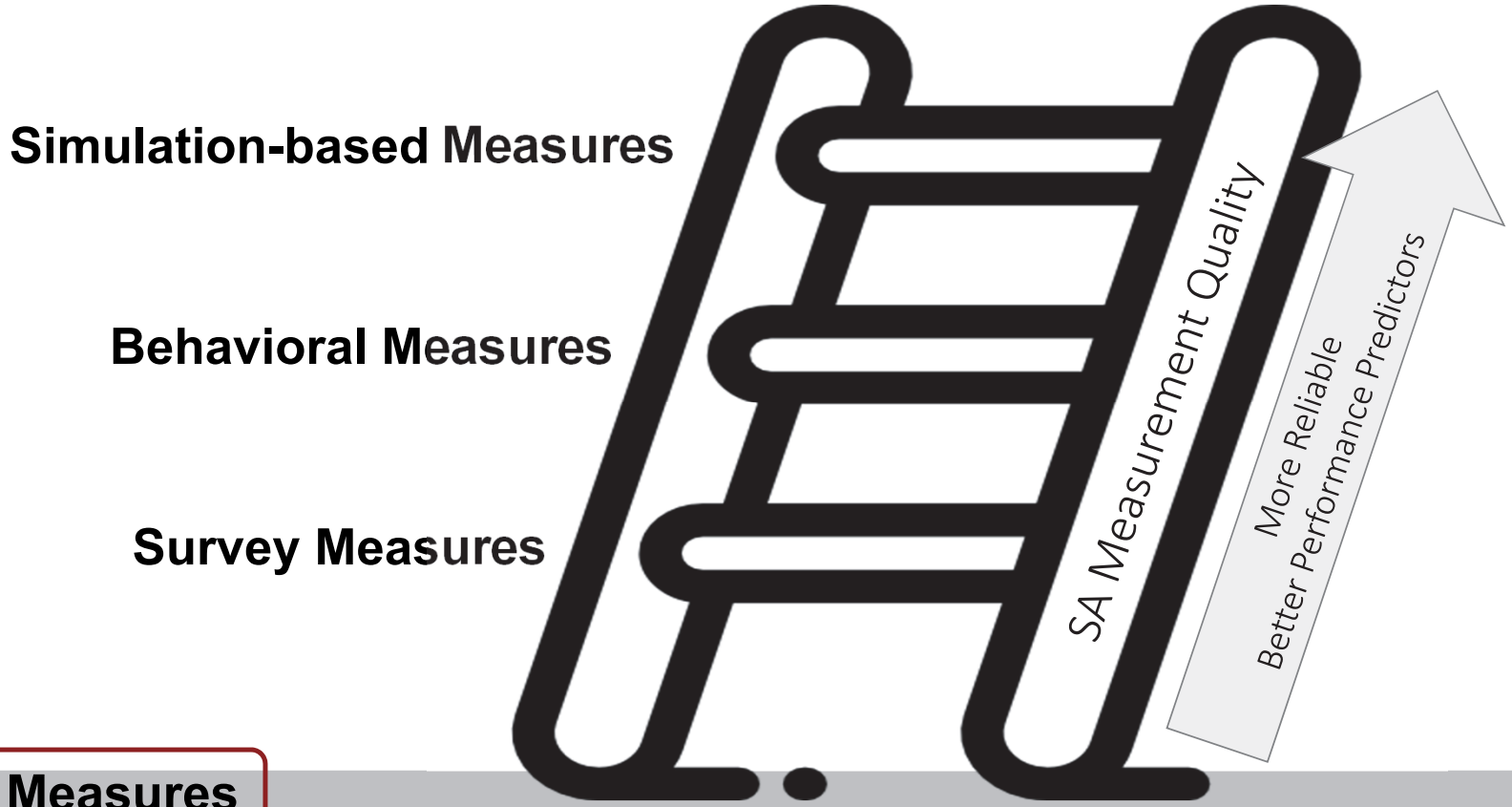
- Easy to implement

Cons:

- Impossible to answer accurately
- Do not account for 3 levels of SA
- Unlikely to reflect ground-truth

OT – Operational Testing; SA – Situation Awareness

Academic Measures of Situation Awareness



Our Current Measures

SA – Situation Awareness

Survey Measures of SA

Survey Scales > Current Measures

- Tested, more reliable
- Scales: SART, SABARS (we will discuss these later)

Operator Surveys: Operator rates their SA at end of task

- Better surveys built based on specific situation

Pros:

- Easy to implement
- Non-intrusive
- Able to compare

Cons:

- Trade-off: ease of use vs. quality
- Subjective measures (*perceptions of* SA, not SA itself)
- Rely on memory

SA – Situation Awareness; SART - Situation Awareness Rating Technique

Behavioral Measures of SA

Operator actions during the mission

- Reflect information operator perceives (level 1 SA)
  - Eye tracking: what operator looks at correlated with what they pay attention to
- Reflect understanding of current state (level 2 SA)
  - Reports problem: description of their understanding
- Reflect anticipation of future events (level 3 SA)
  - Avoidance of obstacles: predicted collision

Pros:

- Real-time measure
- Non-intrusive
- Objective measures

Cons:

- In OT, limited to highly observable behaviors
- Must be tailored to mission, personnel

OT – Operational Testing; SA – Situation Awareness

Simulation-Based Measures of SA

Operators queried when performing a simulated task

- Example: Situation Awareness Global Assessment Technique (SAGAT)
  - Simulation stops at randomly determined intervals
  - Operator queried while displays are blank
  - High accuracy indicates high situation awareness

Pros:

- Objective measures
- Measure all 3 SA levels
- Strongly predict performance

Cons:

- Not conducive to OT
- Must be tailored to mission, personnel

SA – Situation Awareness; OT – Operational Testing

Where do We Go From Here?

We can do better in Operational Testing!

- But, it is going to take some work!
  - Existing research on SA measurement in OT is limited.
  - Developing a measurement methodology is not enough; it has to be broadly understood and adopted.

Do the best you can until you know better. Then when you know better, do better.

- Maya Angelou

OT – Operational Testing; SA – Situation Awareness

The Bare Minimum:
Swap Custom Questions for Vetted SA Survey Measures

Situational Awareness Rating Technique (SART) (Taylor 1990)

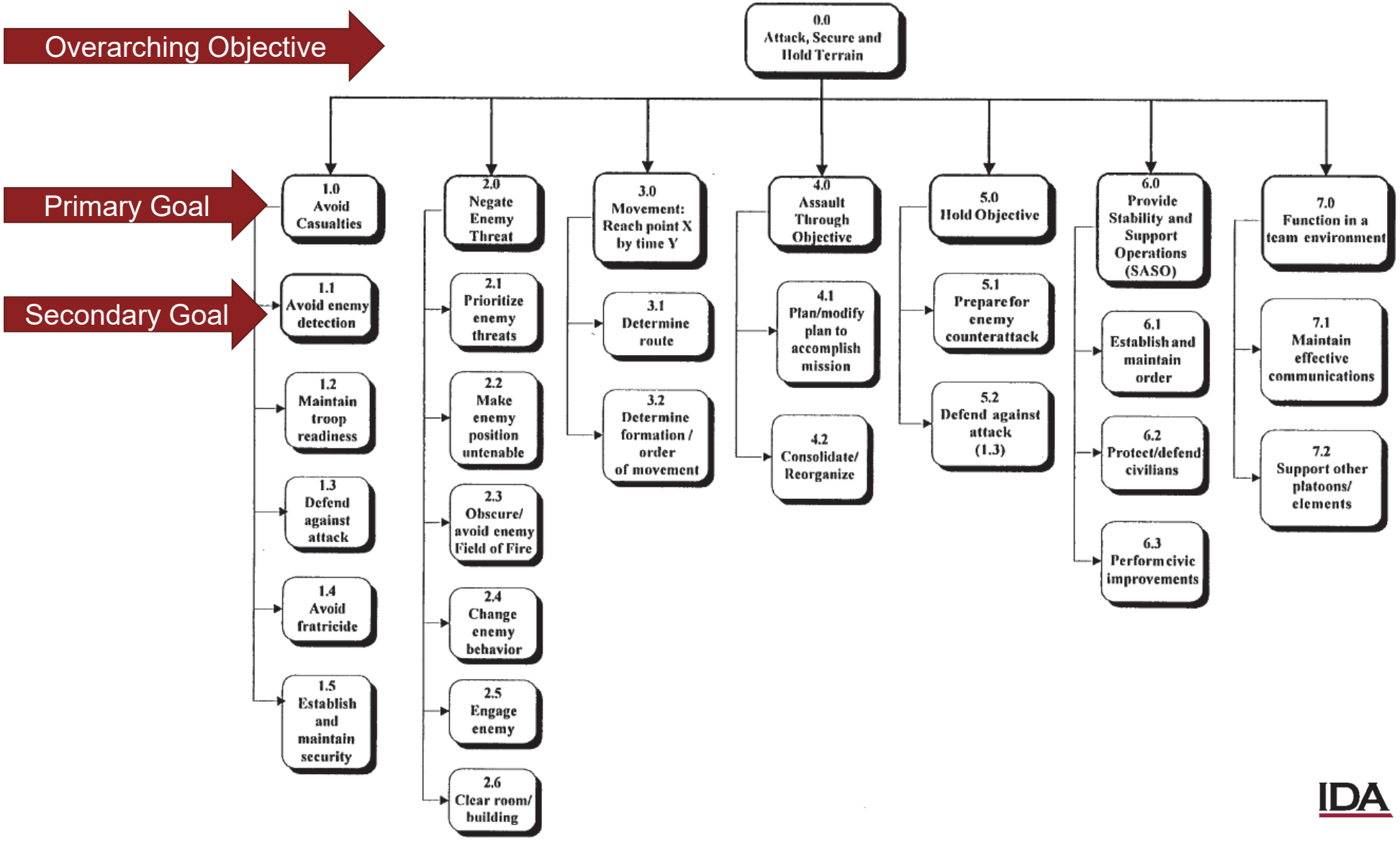
- 10 items
- 3 subscales: attentional demand, attentional supply, and understanding

Except from SART							
How complicated is the situation?							
Simple and straightforward	1	2	3	4	5	6	7 Complex with inter-related components
How much are you concentrating on the situation?							
Focused on one aspect	1	2	3	4	5	6	7 Concentrating on many aspects
How familiar are you with the situation?							
This is a new situation	1	2	3	4	5	6	7 Great deal of relevant experience

SA – situation awareness

CAUTION:

At best, SART measures level 1 SA. We still need to develop better measures


Overarching Objective
Primary Goal
Secondary Goal
0.0 Attack, Secure and Hold Terrain
1.0 Avoid Casualties
1.1 Avoid enemy detection
1.2 Maintain troop readiness
1.3 Defend against attack
1.4 Avoid fratricide
1.5 Establish and maintain security
2.0 Negate Enemy Threat
2.1 Prioritize enemy threats
2.2 Make enemy position untenable
2.3 Obscure/ avoid enemy Field of Fire
2.4 Change enemy behavior
2.5 Engage enemy
2.6 Clear room/ building
3.0 Movement: Reach point X by time Y
3.1 Determine route
3.2 Determine formation / order of movement
4.0 Assault Through Objective
4.1 Plan/modify plan to accomplish mission
4.2 Consolidate/ Reorganize
5.0 Hold Objective
5.1 Prepare for enemy counterattack
5.2 Defend against attack (1.3)
6.0 Provide Stability and Support Operations (SASO)
6.1 Establish and maintain order
6.2 Protect/defend civilians
6.3 Perform civic improvements
7.0 Function in a team environment
7.1 Maintain effective communications
7.2 Support other platoons/ elements

Secondary Goal

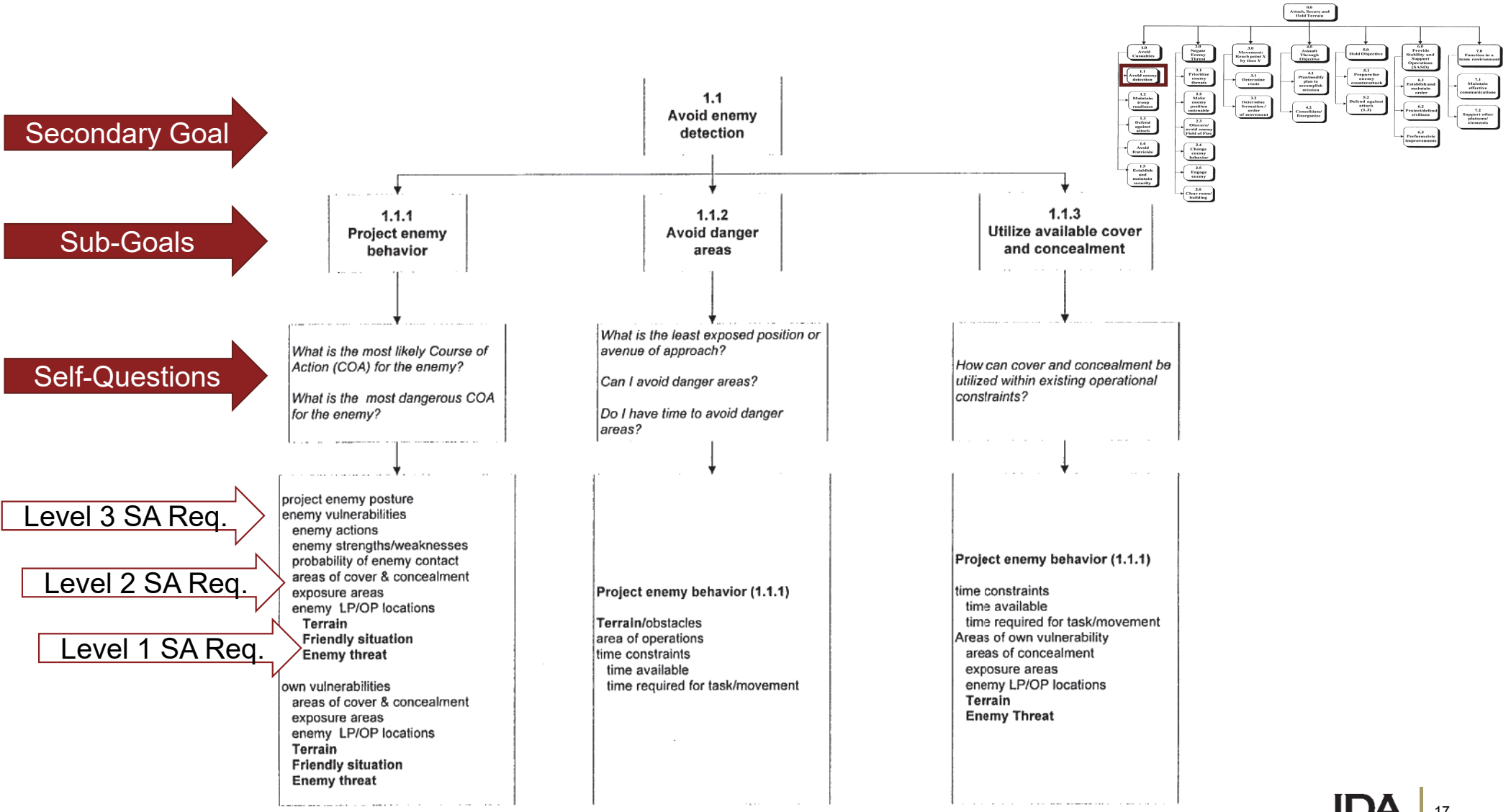
Sub-Goals

Self-Questions

Level 3 SA Req.

Level 2 SA Req.

Level 1 SA Req.



A "Better" Approach: Combine Survey and Behavioral Measures of SA

- Develop a battery of reasonably "easier-to-implement" SA measures. May include:
  - Surveys (e.g., SART) +
  - Behavioral observation (i.e., process indices)
  - SME Ratings (e.g., SABARS)

***Excerpt* from SABARS for Small Unit Leadership in Combat Fatigue Course**	
Information Management	Assigns tasks to squad members based on ability.
	Locates self at vantage point to observe main effort at the objective.
	Utilizes a leader's recon to assess terrain and situation and to finalize plan.
	Uses maps to route find and monitor progress toward objective.
	Maintains appropriate squad security posture throughout mission.
	Selects the appropriate type and amount of equipment and ammunition for the mission.
	Maintains knowledge of time constraints and mission event timing.

SA – Situation Awareness; SME – Subject Matter Expert

Gold-Star Goals:
Excursion as Simulation-Based Measures of SA

- Situation Awareness Global Assessment Technique (SAGAT)
  - Simulation/exercise paused at randomly pre-determined intervals
  - Information about the battlespace is removed. Operators queried during the pause.
  - High accuracy at each level of SA indicates high situation awareness

Example Level 1 Queries	Example Level 2 Queries	Example Level 3 Queries
• Mark the location of each [indicated] element on the map. • How many casualties have your suffered? • Which friendly elements are not in communication with you? • Which rooms in the current building are clear?	• Which enemy location(s) are the weakest/strongest? • Which friendly location(s) are the weakest/strongest? • Which enemy element is your highest threat? • Who has the advantage in the current situation?	• What do you expect the enemy to do in the next 5 minutes? • How long will it take to get to your next position? • Will you be exposed to the enemy on the route to your next position? • Will there be civilians in the line of fire?

SA – Situation Awareness

Questions

- Clarifications on substantive material
- Questions, comments, or concerns
- Next steps / future presentations

References & Resources

Endsley, Mica R. "Design and evaluation for situation awareness enhancement." *Proceedings of the Human Factors Society annual meeting*. Vol. 32. No. 2. Sage CA: Los Angeles, CA: Sage Publications, 1988.

Endsley, Mica R. "Measurement of situation awareness in dynamic systems." *Human factors* 37.1 (1995): 65-84.

Endsley, Mica R., et al. *Modeling and measuring situation awareness in the infantry operational environment*. TRW SYSTEMS INTEGRATION GROUP FAIRFAX VA SYSTEMS DIV, 2000.

Endsley, Mica R. *Situation Awareness Measurement: How to Measure Situation Awareness in Individuals and Teams*. Human Factors and Ergonomics Society, 2021.

Matthews, Michael D., and Scott A. Beal. *Assessing situation awareness in field training exercises*. Military Academy West Point NY Office of Military Psychology and Leadership, 2002.

Matthews, Michael D., Beal, Scott A., and Pleban, Robert J. *Situation awareness in a virtual environment: Description of a subjective assessment scale*. US Army Research Institute for the Behavioral and Social Sciences, 2002.

Matthews, Michael D., et al. "Situation awareness: Predicting small unit leader performance during a combat fatigue course." *Proceedings of the 7th defense analysis seminar*. Washington, DC: Office of the Undersecretary of the Army, 2004.

Salmon, Paul M., et al. "Situation awareness measurement: A review of applicability for C4i environments." *Applied ergonomics* 37.2 (2006): 225-238.

Salmon, Paul M., et al. "Measuring situation awareness in command and control: a comparison of measures study." *Proceedings of the 14th European Conference on Cognitive Ergonomics*. Vol. 250. Association for Computer Machinery, 2007.

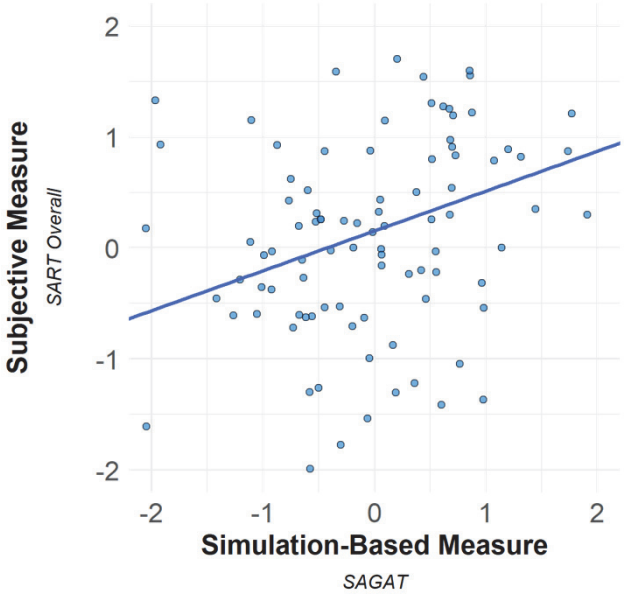
Strater, Laura D., et al. *Measures of platoon leader situation awareness in virtual decision-making exercises*. TRW INC FAIRFAX VA SYSTEMS AND INFORMATION TECHNOLOGY GROUP, 2001.

Taylor, Richard M. "Situational awareness rating technique (SART): The development of a tool for aircrew systems design." *Situational awareness*. Routledge, 2017. 111-128.

Subjective SA Measures Can be Misleading!

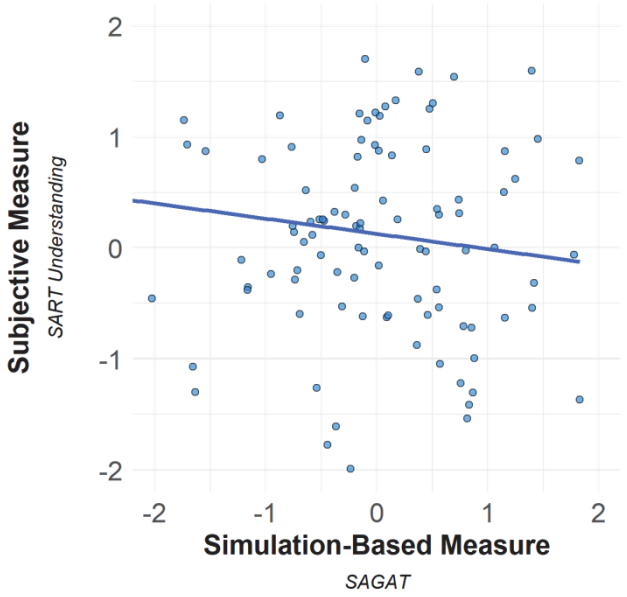
A Comparison of Situation Awareness Measurement Methods in Air Traffic Control

Subjective vs Level 1 SA



r(L1 Perception, SART) = .31

Subjective vs Level 2/3 SA



r(L2/L3 Comp. & Proj., Subjective) = -.33

Comp. – Comprehension; Perc. – Perception; Proj. – Projection; SAGAT – Situation Awareness Global Assessment Technique; SART: Situation Awareness Rating Technique

* Simulated data based on published correlations. Data based on ATCs.

Good SA Assessment Metrics are Objective and Specific to Domain, Mission, System, and User

- Examples of SAGAT Level 1,2, and 3 queries included in simulations:

Level 1	1) Aircraft call sign / altitude / heading 2) Is the aircraft climbing, descending, or level? 3) Is the aircraft in a right turn, left turn, or straight?
Level 2	4) Which pairs of aircraft have lost separation? 5) Which aircraft are currently being impacted by weather?
Level 3	6) What is the next sector aircraft will transition to? 7) Which aircraft will be leaving the sector in the next two minutes? 8) Which aircraft will lose separation if they stayed on their current (assigned) courses? 9) Which aircraft will be impacted by weather in the next five minutes unless action is taken to avoid it?

*(Endsley, Sollenberger, & Stein, 2000)

Yes, this asks for a lot of information from operators (here, ATCs).

These types of data needs are currently going unanswered during T&E, including OT&E.

When planned data collection is inadequate, analyses are also inadequate.

- ATCs – Air Traffic Controllers; SA – Situation Awareness; SAGAT – Situation Awareness Global Assessment Technique

Our Plan for the Next 1–3 Fiscal Years

- Goals:
  - Further investigate SA measurement methods being used.
  - Develop products to promote a scientific understanding of what SA *is* and *is not.*
  - Develop an SA measurement tool by drawing from existing methodologies.
  - Develop examples that can be used for training and as models for future test plan development.

SA – Situation Awareness

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
XX-04-2023	Final	START DATE	END DATE Apr 2023

4. TITLE AND SUBTITLE
DATAWorks 2023 Situation Awareness Should be Measured Scientifically: Progress Towards Evaluation Practices for Operational Testing

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
HQ0034-19-D-0001		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
BD-09-2299	2299(90)	

6. AUTHOR(S)
Armstrong, Miriam, E.; Green, Elizabeth, A.; Mantua, Janna, R.; Vickers, Brian, D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Institute for Defense Analyses 730 East Glebe Road Alexandria, Virginia 22305	NS D-33403 H 2023-000053

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER
Director, Operational Test and Evaluation 1700 Defense Pentagon Room 1D548 Washington, DC 20301	DOT&E	

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES
Project Leader: Haman, John T.

14. ABSTRACT
Situation Awareness (SA) plays a key role in decision making and human performance; higher operator SA is associated with increased operator performance and decreased operator errors. In the most general terms, SA can be thought of as an individual's "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future."

While "situational awareness" is a common suitability parameter for systems under test, there is no standardized method or metric for quantifying SA in operational testing (OT). This leads to varied and suboptimal treatments of SA across programs and test events. Current measures of SA are exclusively subjective and paint an inadequate picture.

Future advances in system connectedness and mission complexity will exacerbate the problem. We believe that technological improvements will necessitate increases in the complexity of the warfighters' mission, including changes to team structures (e.g., integrating human teams with human-machine teams), command and control (C2) processes (e.g., expanding C2 frameworks toward joint all-domain C2), and battlespaces (e.g., overcoming integration challenges for multi-domain operations). Operational complexity increases the information needed for warfighters to maintain high SA, and assessing SA will become increasingly important and difficult to accomplish.

IDA's Test science team has proposed a piecewise approach to improve the measurement of situation awareness in operational evaluations. The aim of this presentation is to promote a scientific understanding of what SA is (and is not) and encourage discussion amongst practitioners tackling this challenging problem. We will briefly introduce Endsley's Model of SA, review the trade-offs involved in some existing measures of SA, and discuss a selection of potential ways in which SA measurement during OT may be improved.

15. SUBJECT TERMS
situation awareness; situational awareness; Human Systems Interactions (HSI); human systems interactions measures

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES\
a. REPORT	b. ABSTRACT	c. THIS PAGE	SAR	36
Unclassified	Unclassified	Unclassified		

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER
John T. Haman	703-845-2132

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE.
Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE.
State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED.
Indicate the time during which the work was performed and the report was written.

4. TITLE.
Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER.
Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER.
Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER.
Enter all program element numbers as they appear in the report, e.g. 61101A

5d. PROJECT NUMBER.
Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER.
Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER.
Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S).
Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S).
Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S).
Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT.
Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.