INSTITUTE FOR DEFENSE ANALYSES

2022 ATEAS Workshop: Evaluating HSI with AI-Enabled Systems: What should you consider in a TEMP?

Rachel Haga, Project Leader

Brian Vickers
Daniel Porter

April 2022

IDA Document NS-D-33067

Log: H 2022-000165

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 BC-9-4908.0-2, "Joint Artificial Intelligence Center" 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 Mr. Vincent "Bram" Lillard and consisted of «Reviewers» from the Operational Evaluation Division.

For more information:
Rachel Haga, Project Leader
rhaga@ida.org •703-845-2768

Vincent "Bram" Lillard, Director, Operational Evaluation Division
vlillard@ida.org • (703) 845-2830

Rigorous Analysis | Trusted Expertise | Service to the Nation

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-33067

2022 ATEAS Workshop: Evaluating HSI with AI-Enabled Systems: What should you consider in a TEMP?

Rachel Haga, Project Leader

Brian Vickers
Daniel Porter

Executive Summary

The Institute for Defense Analyses and the Joint Artificial Intelligence Center (JAIC) are developing a framework to support testers as they evaluate Human Systems Interaction (HSI) with AI-enabled systems. We are creating this framework because HSI expertise in the Department of Defense Test and Evaluation workforce is rare but increasingly important with more automation in systems. Many test teams do not have an HSI expert but they will still have to test and evaluate a system's HSI characteristics. Part 1 of the framework is a practical tool to empower testers when writing and assessing Test and Evaluation Master Plans (TEMPs) for AI-enabled systems.

This short briefing overviews the core areas of Part 1 and conveys a sense of what the fuller framework looks like. More specifically, this briefing focuses on core HSI concerns relevant in a TEMP, covering the following subset of HSI topics in the framework:

- A warfighter's **mental model** is their knowledge about how the system functions. Mental models allow users to predict how the system will function under various conditions.
- A warfighter's **trust** of the system is how much they feel they can depend on it in a vulnerable or uncertain situation.
- A warfighter's **workload** is how many resources they have available compared with what the task requires. Systems must have an appropriate level of workload in the context of accomplishing the mission during operations.
- **Function allocation** is the assignment of taskwork between the warfighter and system. We must make sure this allocation is reasonable and achievable.

This briefing ends with an example of how increased autonomy in aviation contributed to the loss of Air France Flight 447, costing 228 people their lives. Clearly, HSI will continue to be vital in the creation and adoption of effective, suitable, and trustworthy AI-enabled systems.

Evaluating HSI with AI-Enabled Systems: *What should you consider in a TEMP?*

Rachel Haga, Daniel Porter, Brian Vickers – Institute for Defense Analyses

In support of Dr. Jane Pinelis, Chief, AI Assurance

Joint Artificial Intelligence Center (JAIC)

ATEAS 2022

4/14/2022

Attachment

Offloading warfighter work to AI doesn't make HSI less important.

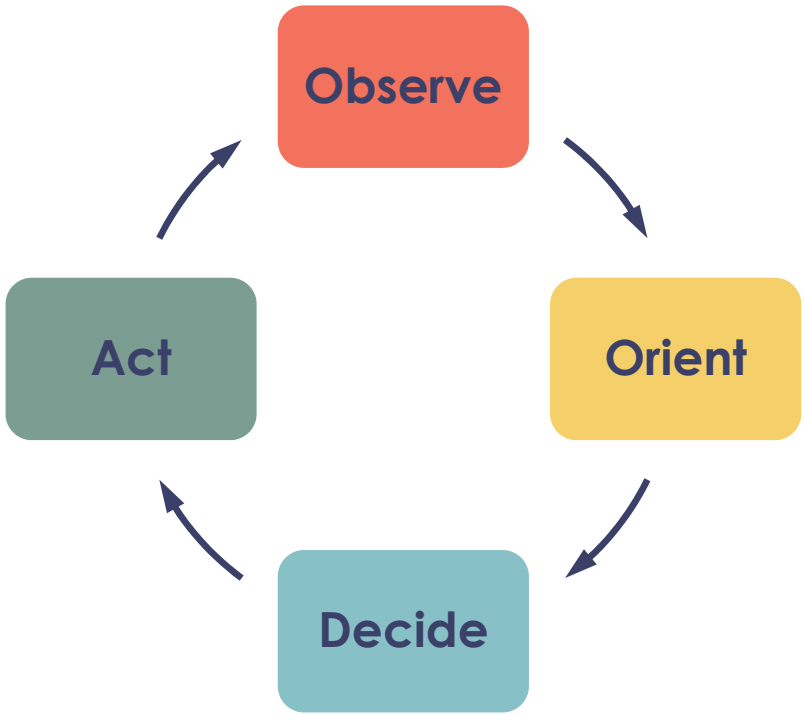
It makes it more important than ever.

What are we talking about today?

1. HSI-AI Framework Overview
2. Some Topics in the HSI-AI Framework

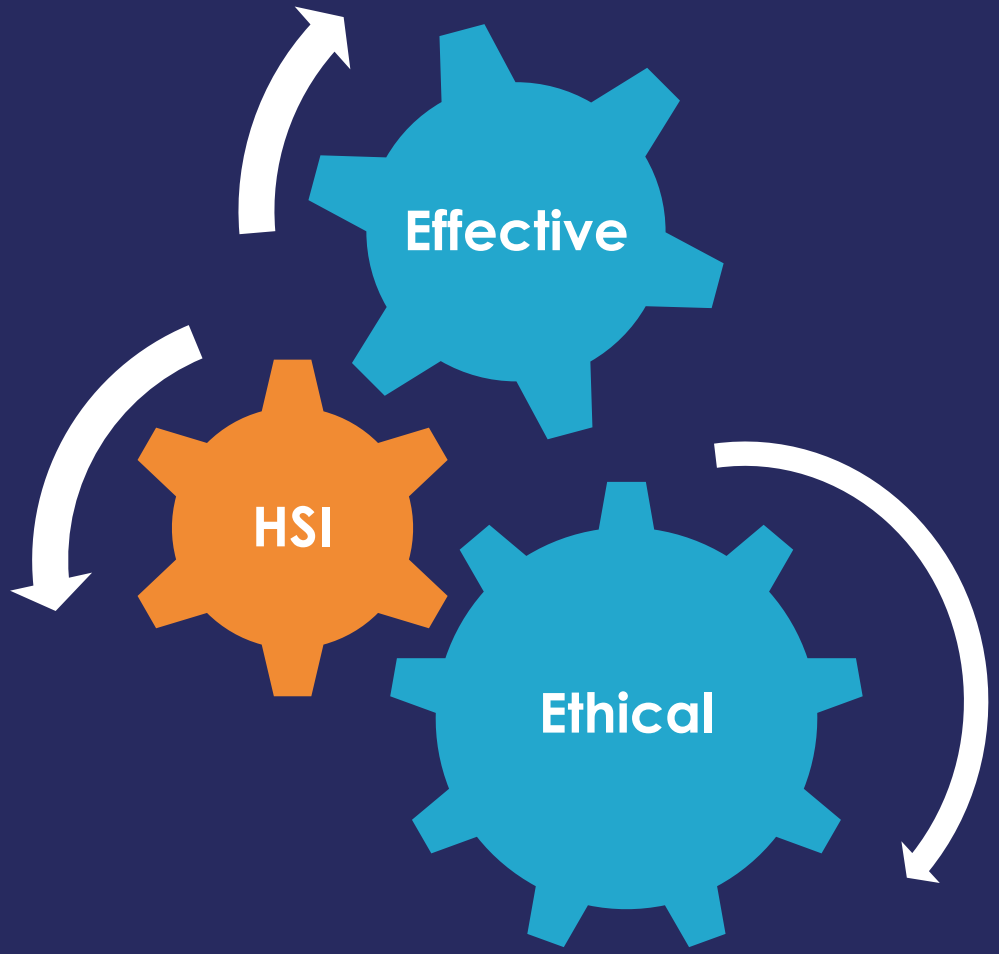
   Mental Models, Workload, Trust, and Function Allocation
3. Current HSI with Complex Systems
4. Takeaways

HSI is about preserving and enhancing performance



O O	"I need to understand and predict the situation. I need what I need to know, when I need to know it, in a way that I understand."
D	"I need to be able to make good decisions about where and how to use this system."
A	"I need to be able to get the system to do what I need and intend it to do."

What you need for effective employment is also what you need for ethical employment



Quality HSI drives both effective and ethical employment

O **O** "I need to understand and predict the situation. I need what I need to know, when I need to know it, in a way that I understand."

D "I need to be able to make good decisions about where and how to use this system."

A "I need to be able to get the system to do what I need and intend it to do."

DoD Ethical AI Principles

- Responsible
- Equitable
- Traceable
- Reliable
- Governable

Framework focuses on working-level testers' tasks

<2% of T&E workforce with any HSI-relevant background

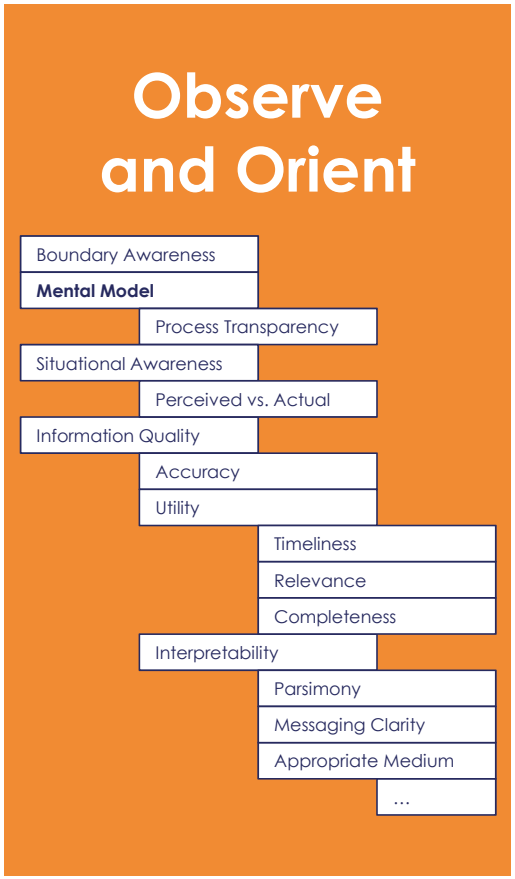
Need to empower non-experts to get 80-90% solution

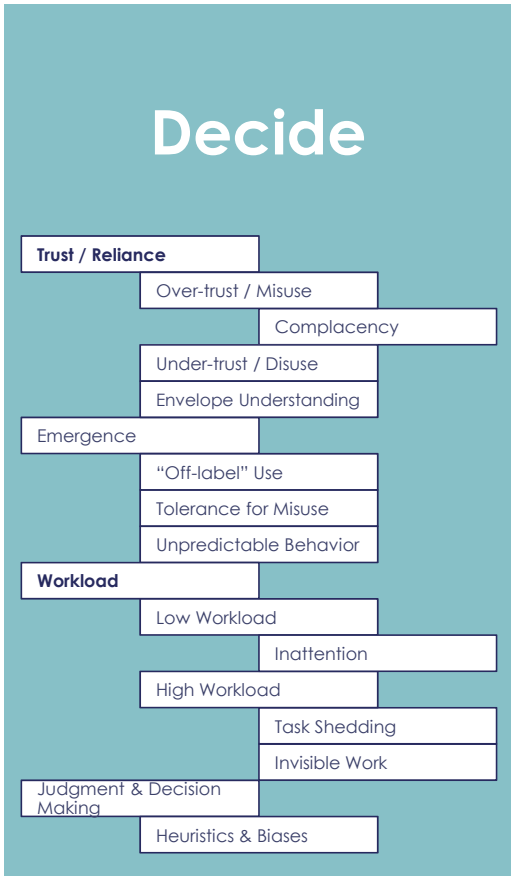
TEMPS
Educate testers about concepts and AI-specific concerns

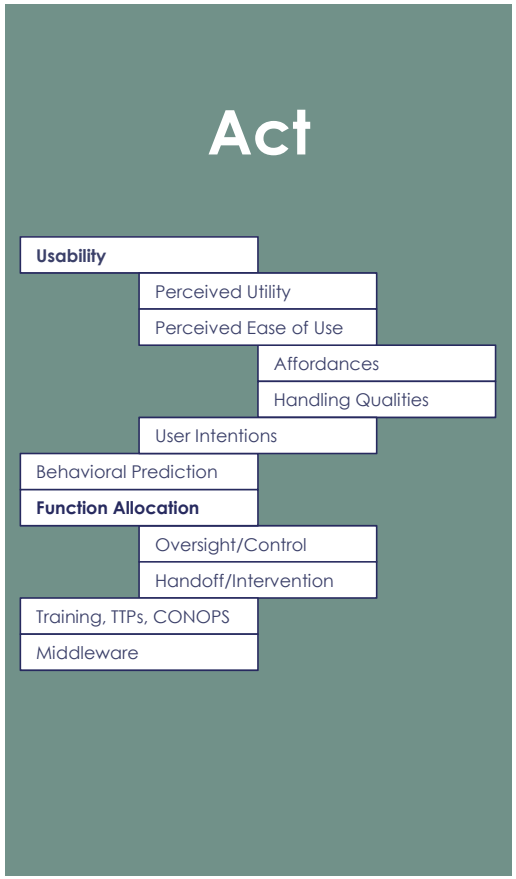
Test Plans
Explain existing techniques, develop new ones

Negotiating
Empower testers by implementing best practices as policy

Products
Provide repos, widgets, and guides for common tasks

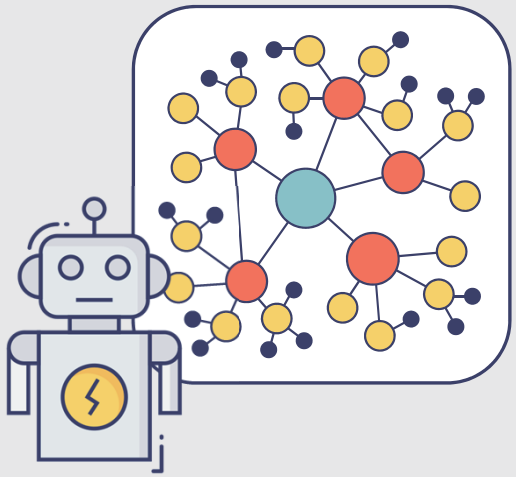

Observe and Orient
Boundary Awareness
Mental Model
Process Transparency
Situational Awareness
Perceived vs. Actual
Information Quality
Accuracy
Utility
Timeliness
Relevance
Completeness
Interpretability
Parsimony
Messaging Clarity
Appropriate Medium
...


Decide
Trust / Reliance
Over-trust / Misuse
Complacency
Under-trust / Disuse
Envelope Understanding
Emergence
"Off-label" Use
Tolerance for Misuse
Unpredictable Behavior
Workload
Low Workload
Inattention
High Workload
Task Shedding
Invisible Work
Judgment & Decision Making
Heuristics & Biases

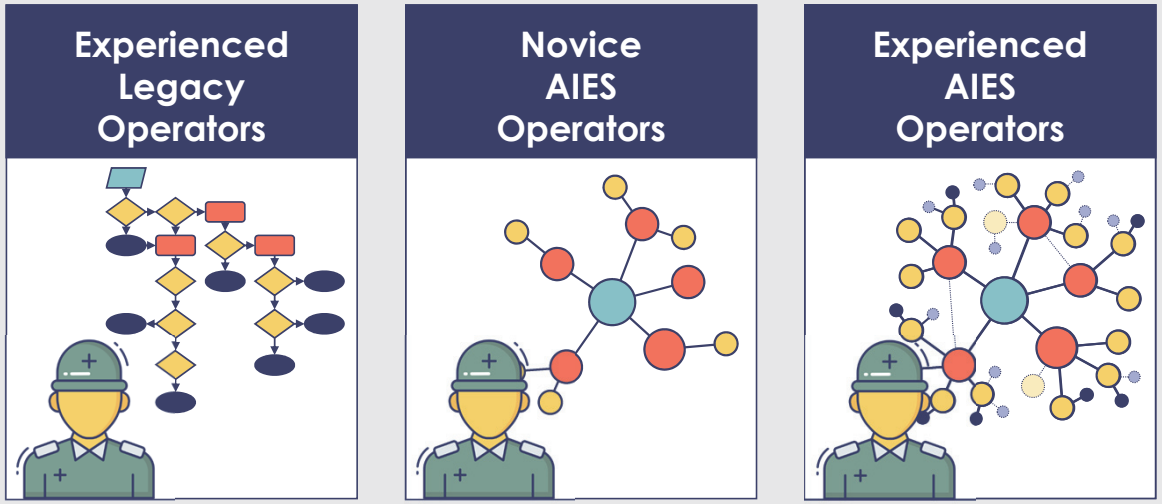

Act
Usability
Perceived Utility
Perceived Ease of Use
Affordances
Handling Qualities
User Intentions
Behavioral Prediction
Function Allocation
Oversight/Control
Handoff/Intervention
Training, TTPs, CONOPS
Middleware

Building Mental Models

Humans form **mental models** of automation that allow them to infer the current state of a system from incomplete information and make predictions about future states based on the current one.



System's underlying model



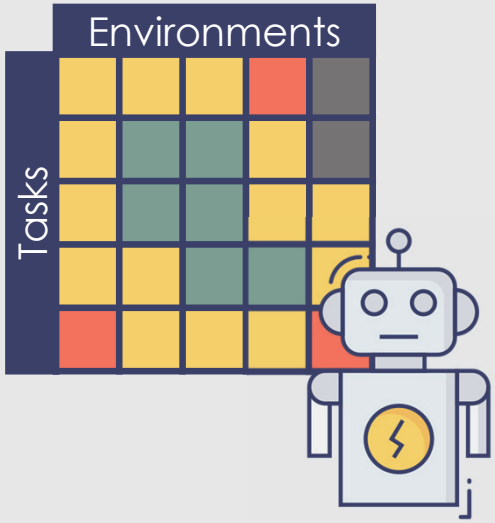
Warfighter's mental model of system behavior

Calibrating Trust

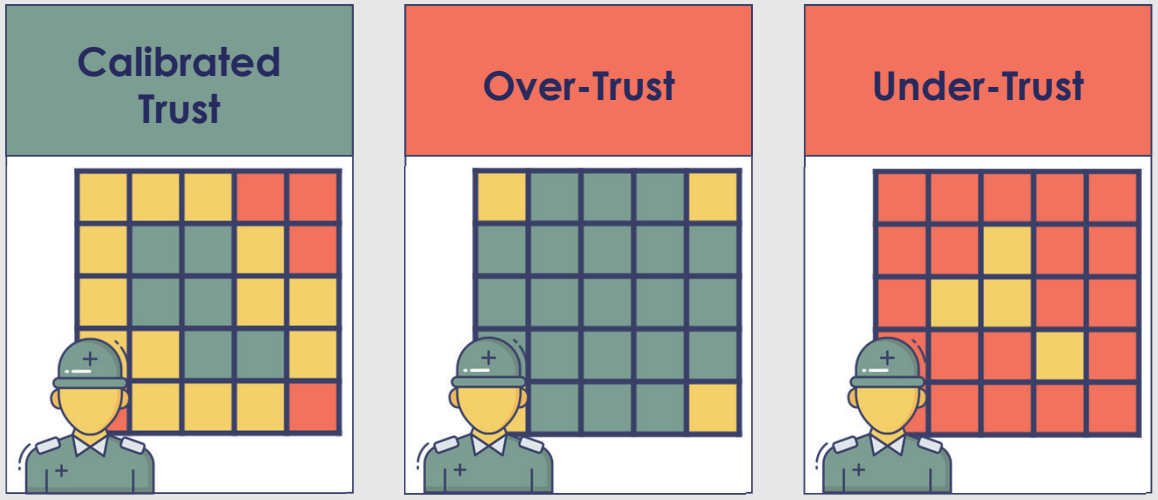
Trust is a person's belief that something can be depended on in vulnerable or uncertain situations. The critical outcome of trust is **reliance**, which is the behavioral, continued use of a system.

Calibrating Trust

Calibrated Trust occurs when the warfighter's **operational reliance** aligns with the **system performance** for a given context.



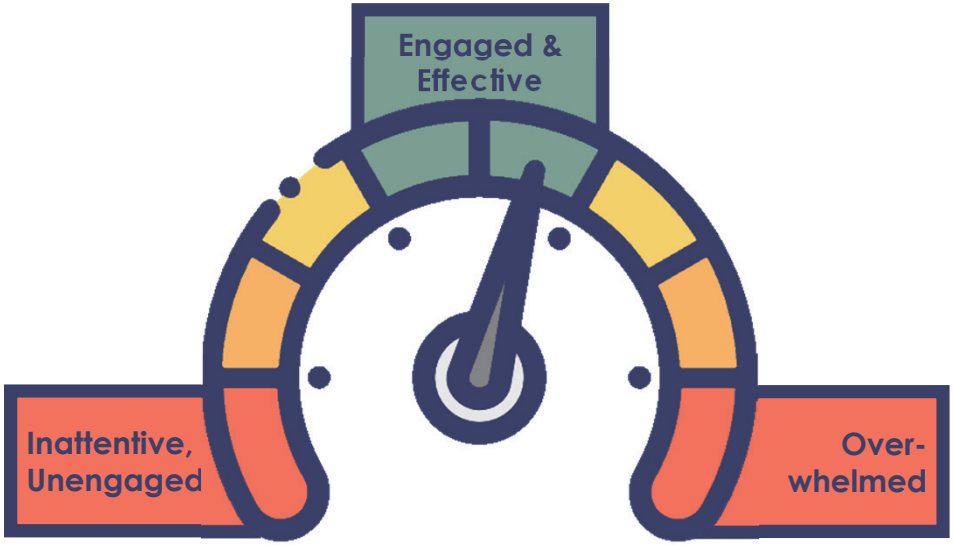
System performance



Warfighter's trust and subsequent reliance on the system

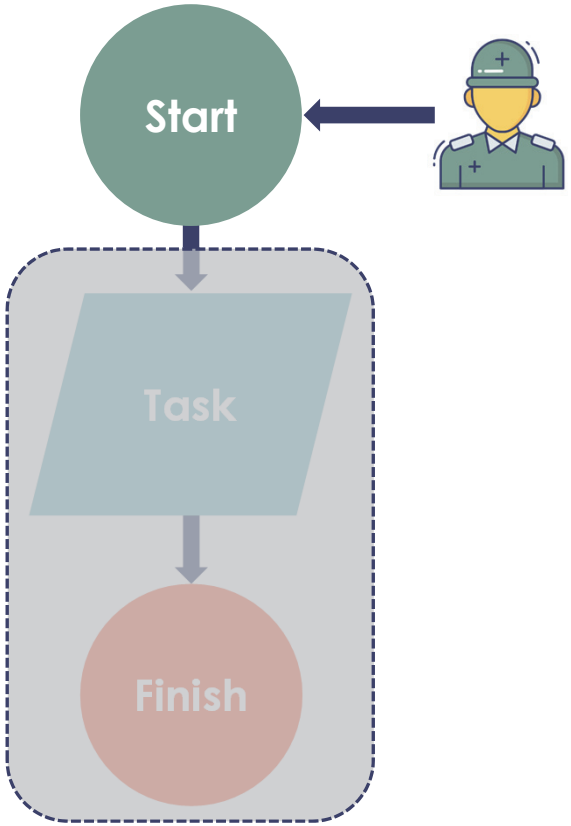
Calibrating Workload

Workload comprises the physical, mental, and temporal resources required by the current tasks relative to the resources available to the person.



Assessing Usability

Usability is the fitness of a tool for a task. Composed of **utility** and **ease of use.**

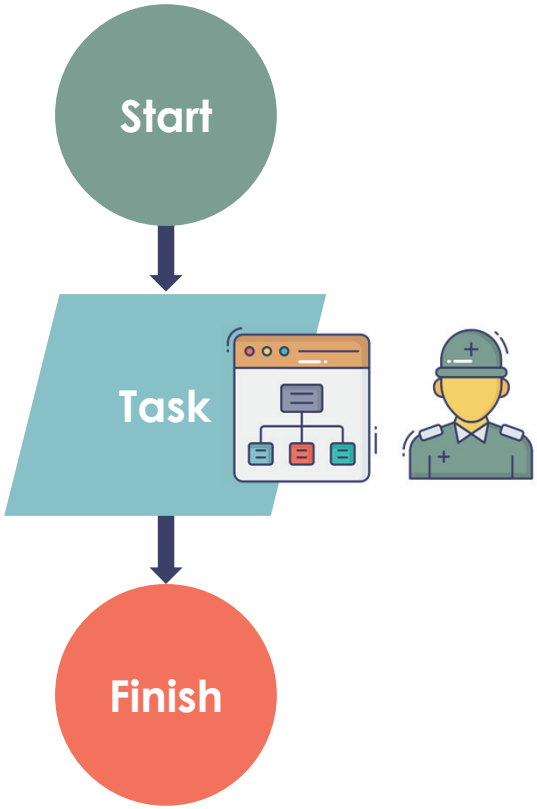


Give Initial Orders

Must assess whether the system achieved the warfighter's intent

Assessing Usability

Usability is the fitness of a tool for a task. Composed of **utility** and **ease of use.**



Give Initial Orders

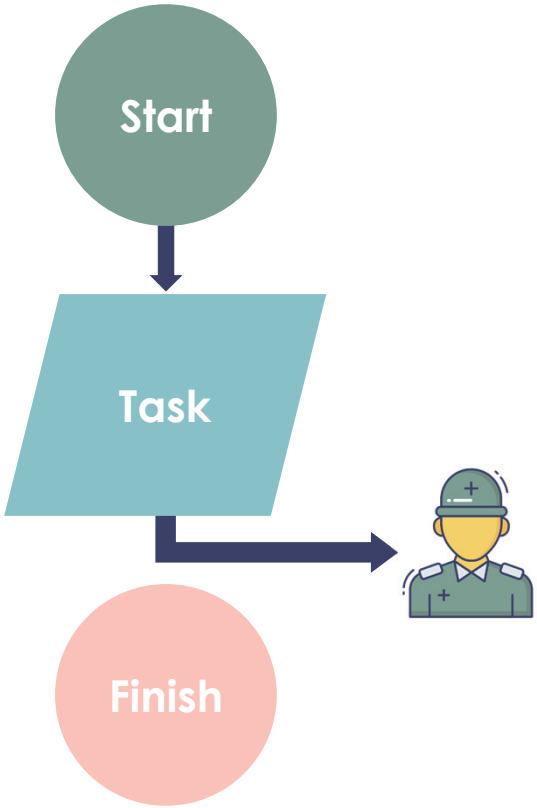
Must assess whether the system achieved the warfighter's intent

Extract Information

Must confirm warfighter has information needed to be situationally aware

Assessing Usability

Usability is the fitness of a tool for a task. Composed of **utility** and **ease of use.**



Give Initial Orders

Must assess whether the system achieved the warfighter's intent

Extract Information

Must confirm warfighter has information needed to be situationally aware

Intervene & Take Over

Must evaluate off-nominal situations where tasks are handed off to warfighter

Assigning Function Allocation

Function allocation is the assignment of the collective taskwork between human and automated agents (including both AIES and traditional software tools) required to achieve mission goals.

Autonomy

Which agent is capable of completing the task?

Authority

Which agent is assigned to a task?

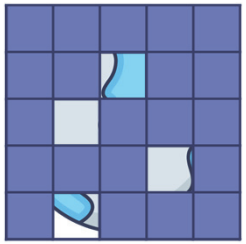
Responsibility

Which agent is accountable for the outcome?

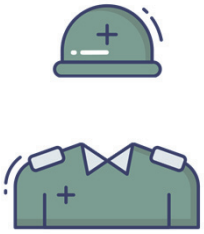
Assigning Function Allocation

Function allocation is the assignment of the collective taskwork between human and automated agents (including both AIES and traditional software tools) required to achieve mission goals.

Confirm that all agents are **capable** of completing their assigned tasks



Confirm that taskwork is intentionally assigned, and that the warfighter is not given the "**leftover allocation**"



Confirm that all "**invisible work**" is accounted for in the CONOPS and test scenarios.

1 2 3 4

Airspeed sensors blocked

Autopilot, stall protections terminated

Unexpected high workload spike

Over-trust has led to complacency

Under-developed mental model

The automation does not display the error

Pilots uncertain of function allocation

Pilots fail to diagnosis problem

On June 1, 2009, a completely functional Airbus A330 crashed into the Atlantic Ocean, killing all 228 people on board.

Airspeed sensors blocked

Autopilot, stall protections terminated

Unexpected high workload spike

Over-trust has led to complacency

Under-developed mental model

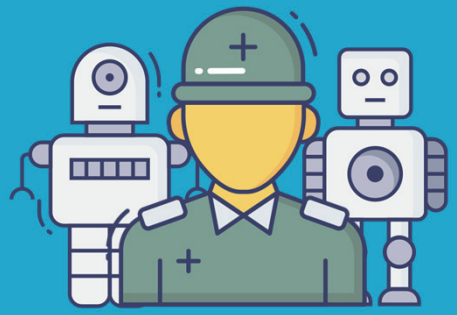
The automation does not display the error

Pilots uncertain of function allocation

Pilots fail to diagnosis problem

Poor HSI can cost people their lives

But we can learn from the past



In a world of AI and Autonomy, the human is still relevant

This has all happened before and it can happen again

The JAIC is putting together an HSI-AI Framework to identify issues

Questions?

Connect with us:
AI.mil
Twitter: @DoDJAIC
LinkedIn: DoD Joint Artificial Intelligence Center

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01 April 2022	OED Draft	N/A

4. TITLE AND SUBTITLE

2022 ATEAS Workshop: Evaluating HSI with AI-Enabled Systems: What should you consider in a TEMP?

5a. CONTRACT NUMBER

HQ0034-19-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rachel Haga
Brian Vickers
Daniel Porter

5d. PROJECT NUMBER

BC-9

5e. TASK NUMBER

4908.0 - 2

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

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

8. PERFORMING ORGANIZATION REPORT NUMBER

D-33697

H 2022-000165

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Joint Artificial Intelligence Center

10. SPONSOR/MONITOR'S ACRONYM(S)

JAIC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

Project Leader: Rachel Haga

14. ABSTRACT

The Institute for Defense Analyses and the Joint AI Center (JAIC) are developing a framework to support testers as they evaluate Human Systems Integration (HSI) with AI-enabled systems. This product is a short briefing that gives an overview of the framework, conveys a sense of what the fuller framework looks like, and ends with an example of how increased autonomy in aviation led to the loss of Air France 447. We are creating this framework because HSI expertise in the DOD Test & Evaluation workforce is rare and increasingly important with more automation in systems. This briefing focuses on the portion of the framework which introduces non-experts to core HSI concerns to help testers write and assess TEMPs. This briefing covers a subset of HSI topics included in the framework, including mental models, trust, workload, and function allocation. HSI will continue to be a vital component of creating and adopting effective, suitable, and trustworthy AI-enabled systems.

15. SUBJECT TERMS

Artificial intelligence (AI)
Human systems Interactions (HSI)
Joint Artificial Intelligence (AI) Center (JAIC)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Rachel Haga
Unclassified	Unclassified	Unclassified	Unlimited	29	19b. TELEPHONE NUMBER *(Include area code)* 703-578-2768