SAND2010-6515C

Sandia National Laboratories

Performance Assessment to Enhance Training Effectiveness

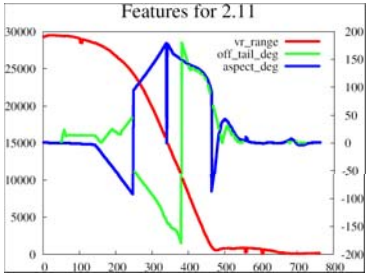
Susan M. Stevens-Adams, Justin D. Basilico, Robert G. Abbott, Charles J. Gieseler & Chris Forsythe

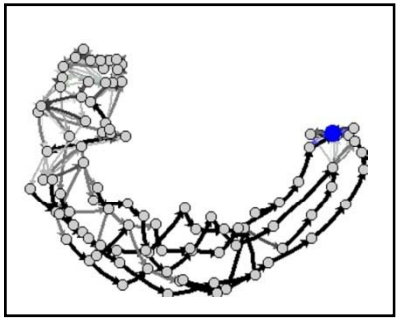
The U.S. military needs technologies enabling instructors to accomplish more with available time and resources.

- *The NAVAIR PMA205 Capability/Technology Gaps Assessment for aviation training systems found a general need for enhancements to improve the effectiveness and efficiency of training operations.*
  - Specific needs included technical innovations for brief-debrief systems, and human-systems improvements for exercise workload reduction and enhanced instructor workstations.
- *The Naval Aviation Simulation Master Plan calls for "unique capability to assess mission execution during post-event debrief…" including needs for "instructor workstations, robust mission playback and debrief capability and the capability to track metrics of aircrew performance."*

The AEMASE Approach – Automated Expert Modeling and Student Evaluation

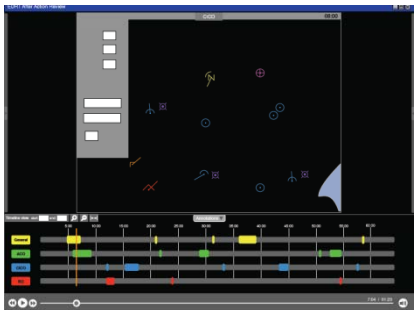
1. Subject matter experts demonstrate desired behavior in a simulator or instrumented environment.



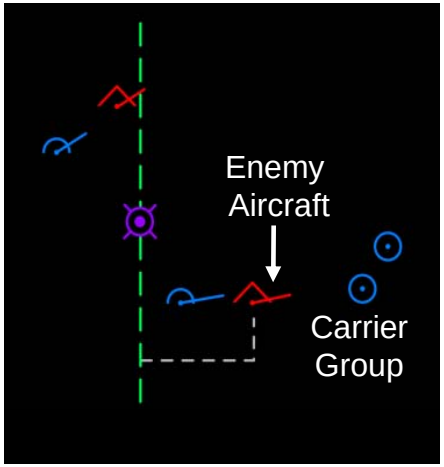


2. Machine learning techniques used to construct a model of expert behavior.

3. During training, student behavior is compared to expert model to identify and target training to individual deficiencies.



AEMASE encompasses representation and assessment of performance metrics

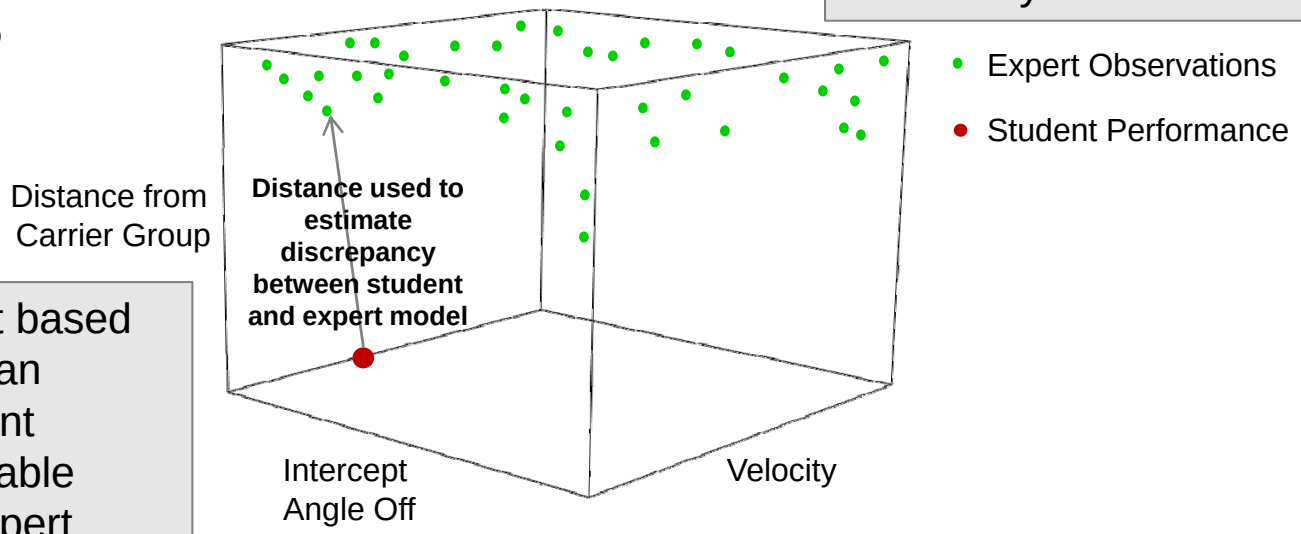


Snapshot from simulation scenario

Various parameters (e.g. distance, angle, velocity) serve as features and are combined to create a vector describing a situation (e.g. relationship between entities)

Feature 1	Feature 2	Feature 3	
Distance from Carrier Group ;	Intercept Angle Off ;	Velocity	Vector

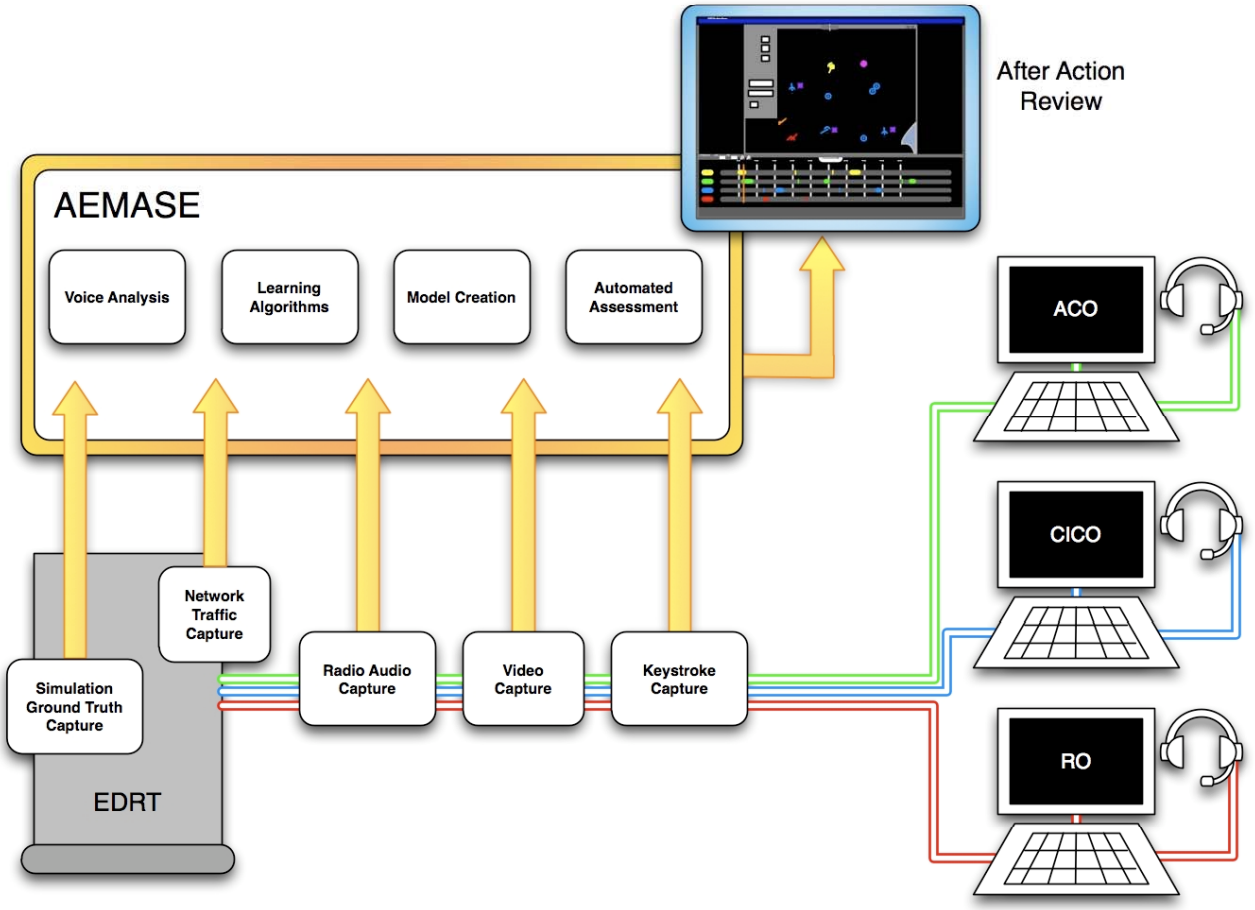
Vectors are treated as points within a multidimensional space defined by the features



Performance assessment based on the distance between an observed point (i.e. student performance) and acceptable points defined through expert performance (i.e. expert observations)

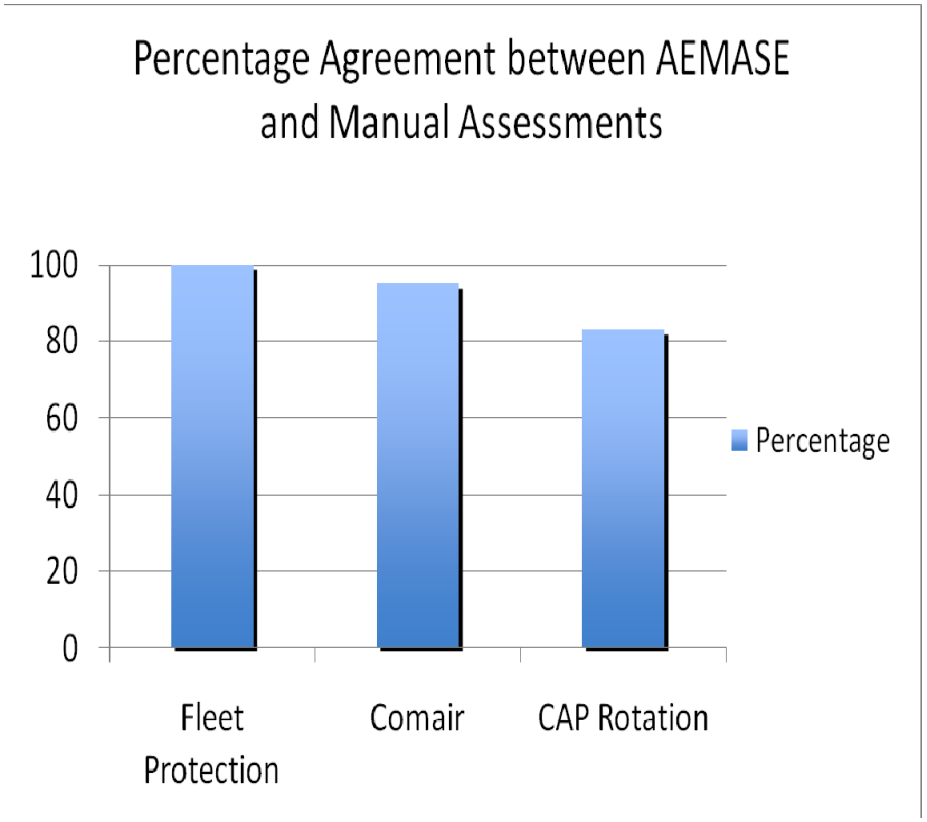
AEMASE capabilities have been integrated with the E-2 Enhanced Deployable Readiness Trainer (E2EDRT)

E2EDRT is an operational trainer deployed at NSAWC Fallon, NAS Point Mugu, NS Norfolk & NAS Atsugi



Previous research demonstrated the accuracy of automated assessments using AEMASE approach

Using illustrative metrics, we have shown good agreement between automated assessments and human raters



Illustrative Metrics

- Fleet Protection = enemy fighter incursions into commit region
- COMAIR = latency to label COMAIR
- CAP Rotation = recognize and respond to gap in air defenses

A primary objective was to evaluate the effectiveness of the AEMASE approach to training

Hypothesis: Will find superior performance for a group in which an instructor debrief tool facilitates training targeted to individual performance deficits, as compared to a group without similar technology assistance.

Subjects – 22 employees of Sandia with demographics matching those of an entry-level E-2 Naval Flight Officer

Experimental Group – 10 subjects

Instructor utilized debrief tool featuring graphical depictions (i.e. timeline & occupancy maps) of student performance

Control Group – 12 subjects

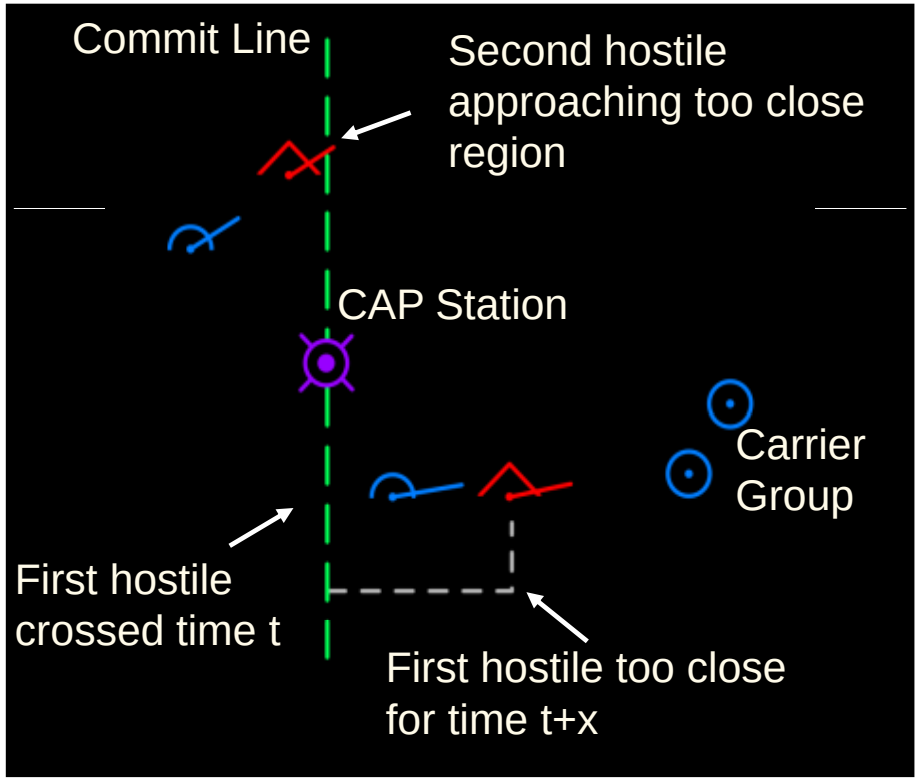
Instructor observed student performance noting and verbally addressing performance deficiencies

Our program of training enabled subjects to complete relatively complex scenarios on the E2EDRT

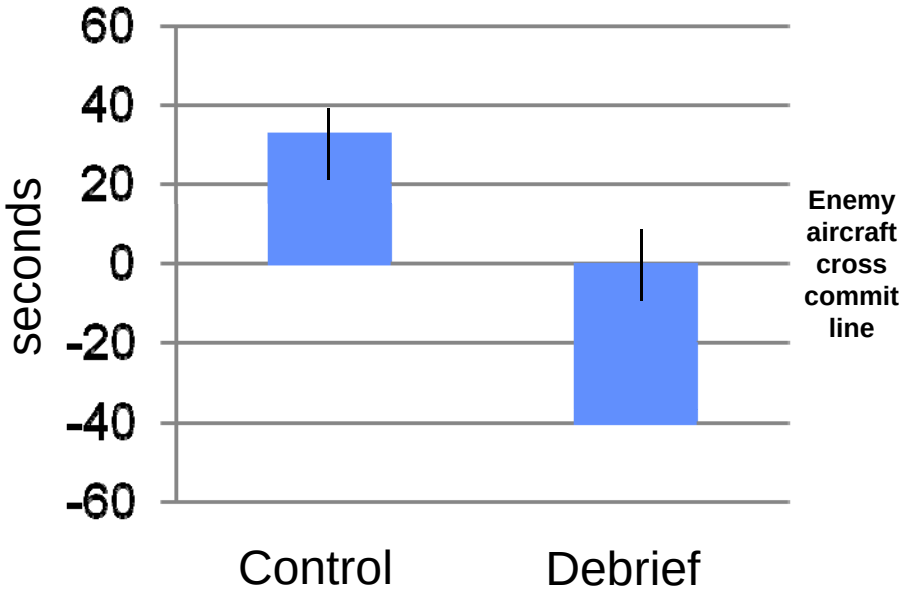
- 8-hr classroom session covering fundamentals
  - e.g. force structure, types of assets, displays & controls, communications, etc.
- 5 simulation-based training sessions
  - E2EDRT Familiarization
    - Objective: familiarization with E2EDRT displays and controls
  - Check-In Procedures and Managing Air Assets
    - Objective: introduce radio communication with AW and familiarize with detect, track and identify air tracks and honoring commit criteria
  - Managing Surface Assets
    - Objective: familiarize student with labeling, identification and management of surface tracks and communication with AZ
  - Tactical Situations I & II
    - Objective: integration of both air and surface pictures in more complex scenarios

Subjects did significantly better recognizing and responding to enemy aircraft threatening the fleet

Student Objective: Prevent non-friendly entities from nearing carrier group (simple metric, but key parameter in assessing NFO performance)



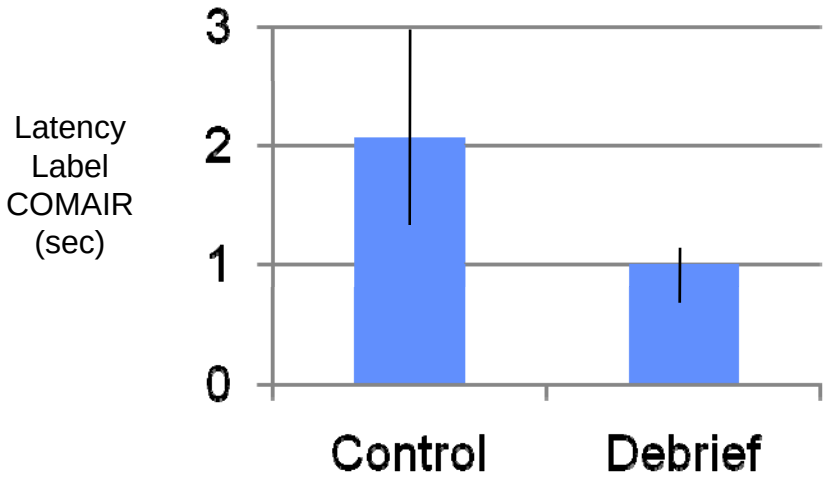
Friendly fighters committed sooner in response to enemy aircraft approaching commit line
(*t* =2.03*; *p<0.05)*



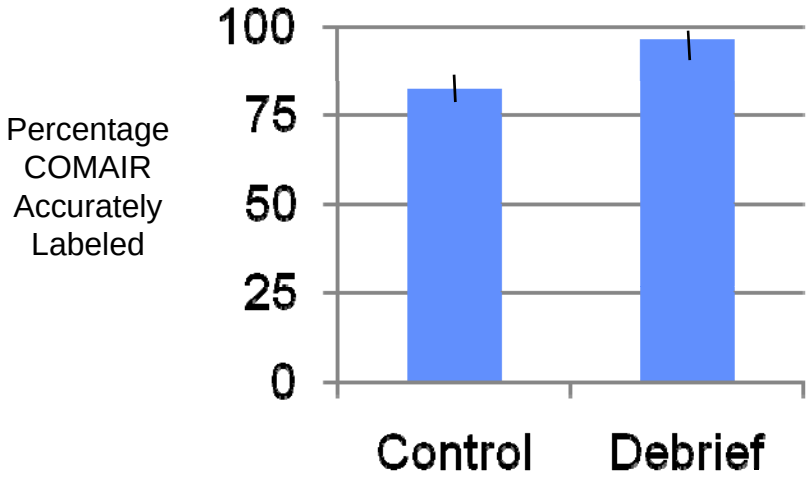
* Results reported for a one-tailed t-test

Significantly superior performance for labeling COMAIR found with debrief tool

COMAIR were labeled significantly faster
(*t* =1.69*; *p<0.05)*



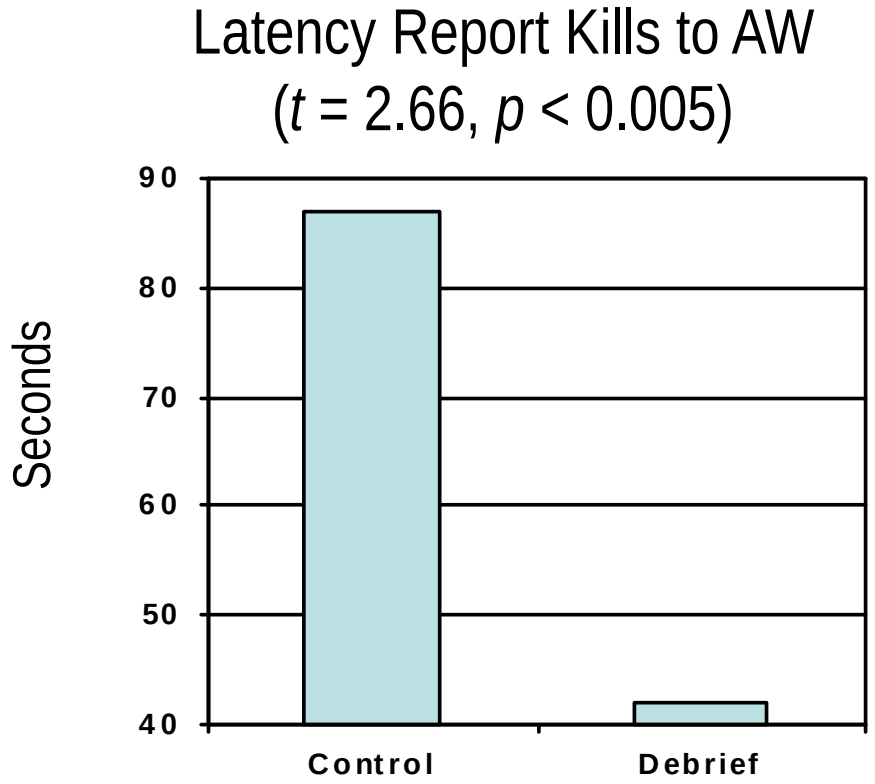
COMAIR labeled significantly more accurately
(*t* = 1.87*; p<0.05)



* Results reported for a one-tailed t-test

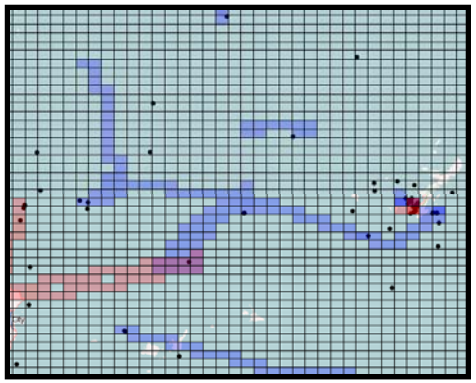
Subjects trained with debrief tool more promptly reported kills to Air Warfare Commander

Student Objective: Once an enemy aircraft is reported downed, E-2 NFO should promptly report to Air Warfare Commander to update battlespace situation awareness

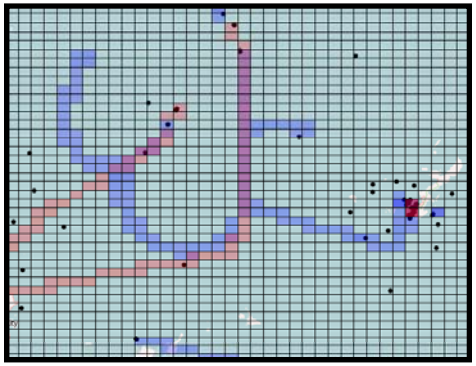


Airspace management proved to be too difficult of concept for the limited training subjects received

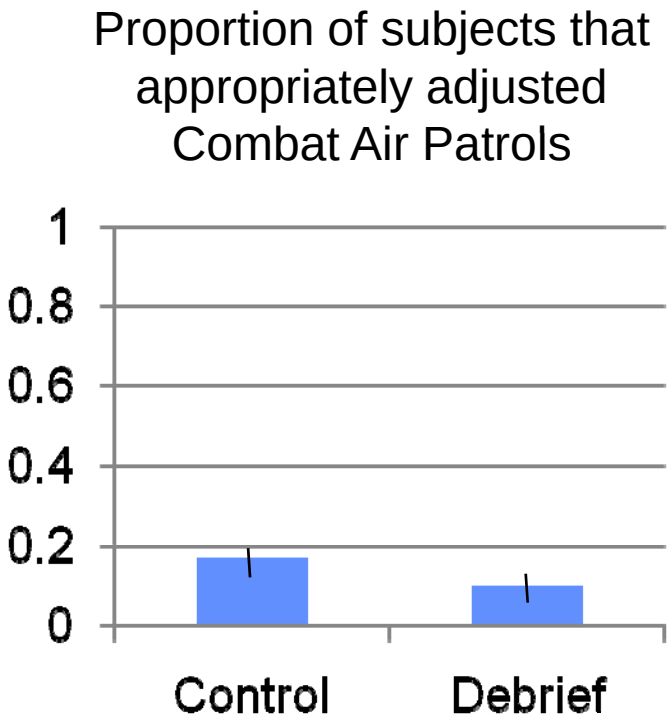
Student Objective: As the battlespace evolves, student should effectively manage their assets, including reassignment in response to developing situations.



Good Tactics - Attack Repelled



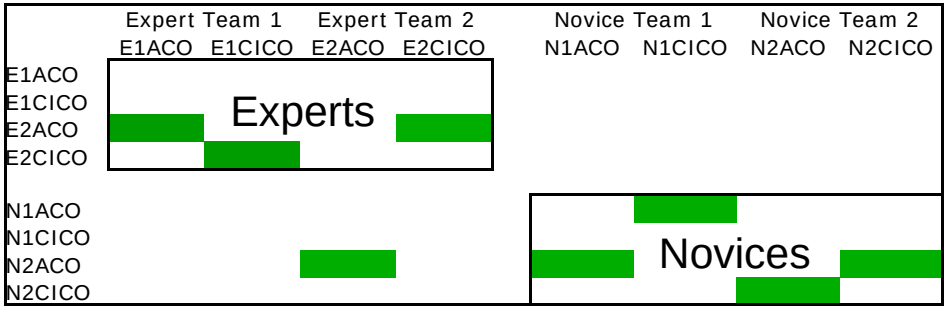
Poor Tactics - Airspace Violated



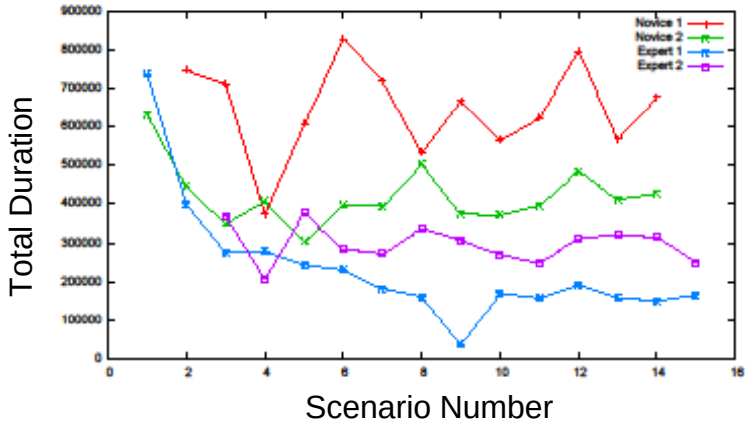
Current efforts are focused on extending automated assessment to verbal communications

Teams consisting of either expert E-2 NFOs or novices performed scenario engineered to stress team coordination

Semantic analysis of radio communications showed experts were more similar to other experts, and novices to other novices



Green indicates who each subject most resembled



The duration of novice radio communications was longer than experts

Novices used more filler words (e.g. “ah,” “er”) than experts

	Experts	Novices
ah	1	6
er	4	8
like	5	9
uh	112	307
um	5	28
Total	127	358

AEMASE approach enables instructors to make more effective use of simulation training technologies

- Use automated assessments to capture mundane events, allowing instructors to focus attention on higher-level knowledge and skills
- Graphical depictions of scenario events facilitate instructors in communicating the “big picture” of what transpired during an exercise
- With automated assessments, there is an opportunity to standardize selected metrics

Acknowledgement

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DEAC04-94AL85000.

Research and development for application of the AEMASE approach to the E-2 was performed through a contract award from the Office of Naval Research.

Data collection with the E2EDRT was partially funded through a contract award from the Defense Advanced Research Projects Agency.