Human Operator Effects in
Theater Ballistic Missile Defense

Year 2001 Multinational BMD Conference and Exhibition
5-8 June 2001
Maastricht, The Netherlands

Dr. Daniel R. Tufano, Sr.
POET/Oak Ridge National Laboratory

Dr. Leslie Pierre
Ballistic Missile Defense Organization

Dr. Charles W. Glover
POET/Oak Ridge National Laboratory

Outline

- **Study Objectives**

- **Experimental Factors (Independent Variables)**

- **Experimental Design Matrices**

- **Subjective Direct Assessment Methods**

- **Detailed Results: Individual Operator Data By Position Played**

- **Future Convergent Methods with Empirical Data**

Study Objectives & Factors

Examine:
- **Centralized vs. Decentralized Human Control as a function of Air Picture Quality**
- **Degrees of Automated Battle Management Aid (ABMA) as a function of Air Picture Quality**

Experimental Factors/Independent Variables

- **Levels of Centralized/Decentralized Human Control**
  - **Physical Deconfliction**
  - **Decentralized with Upper Tier Engagement Coordination**
  - **Centralized Control by the Regional Air Defense Commander (RADC)**

- **Air Picture Quality**
  - **"Perfect"**
  - **Bad: some dual tracks per truth object**
  - **Worst: greatest number of dual tracks per truth object**

- **Degree of Automation/Automated Decision Aiding**
  - **Automated Battle Manager (ABM) for Weapon Target Assignment (WTA)**
  - **Simulated Data Link Message Displaying Shot Opportunities**
  - **Human Control, No Automation, No Shot Opps Information**

Text

Study Objectives & Factors

Objectives: The primary objective was to examine the impact of air/space track picture quality on the execution of Joint Theater Ballistic Missile Defense, when this picture is shared by all players in all positions. Air picture quality impact was studied together with doctrinal variations (I.e., degree of centralized control) and degree of automated battle management aiding.

Factors:

███: In the physical deconfliction case each shooter decided which threats to engage, depending on whether it was heading for an defended asset or area for which that shooter was responsible. Decentralized with upper tier engagement coordination involved the joint assignment of weapons to threats by the Tactical Action Officers and the Tactical Directors responsible for upper tier interceptors. Centralized control by the Regional Air Defense Commander (RADC) required that all weapon target assignments be made by the RADC.

█████████████: The "Perfect" level had no ambiguities in its display of tracks. The "Bad" level showed some dual tracks per truth object. The "Worst" level had the greatest number of dual tracks per truth object.

████████████████████████████████████:   An Automated Battle Manager (ABM) for Weapon Target Assignment (WTA) was used in the fully automated condition. The ABM/WTA used an algorithm which implemented the shot doctrine. There was a window of time during which a human could override the battle manager's WTA. An intermediate level of automated decision aiding involved a simulated data link message displaying shot opportunity information for all threats. Finally, there was a case with human control, no automation, and no shot opps information. All three of these cases followed a centralized doctrine.

Experimental Design

2 Overlapping Factorial Designs: 15 Cells Total

Centralization X Air Picture Quality Effects

Air Picture Quality	Physical Deconfliction	Decentralized W/ UT Coord.	Centralized Control
Perfect			
bad			
worst			

Experimental Effects Addressed:
- **Air Picture Quality Main Effect**
- **Centralization Main Effect**
- **Centralization X Air Picture Quality Interaction**

Automated Battle Management Aiding X Air Picture Quality Effects

Air Picture Quality	Human No Shot Opps Info	Human With Shot Opps Info	Automated Weapon Assignment
Perfect			
bad			
worst			

Experimental Effects Addressed:
- **Air Picture Quality Main Effect**
- **Automated Battle Management Aid Main Effects**
    - **Automated Weapon Target Assignment**
    - **Shot Opportunity Information**
- **Automation X Air Picture Quality Interaction**

Text

Experimental Design

2 Overlapping Factorial Designs: 15 Cells Total

Centralization X Air Picture Quality: This matrix is a 3 X 3 factorial design. There are three doctrinal conditions, increasing in the degree to which command is centralized; and there are three levels of air picture quality. This design supports examination of: 1) An Air Picture Quality Main Effect - Does the quality of the air picture have an impact, regardless of doctrine? 2) Centralization Main Effect - Does the degree of centralized control have an impact, regardless of the quality of the air picture? 3) Centralization X Air Picture Interaction - Are there synergistic effects which emerge only as a function of both factors? Are some levels of centralization more susceptible to degradation of the air picture than others?

Automated Battle Management Aiding X Air Picture Quality: This 3 X 3 factorial design has three conditions of increasing automated aiding, and three levels of air picture quality. All of the test runs used centralized control. The design matrix permits examination of: 1) An Air Picture Quality Main Effect - Does the quality of the air picture have an impact, regardless of the amount of automated aiding? 2) Automated Aiding Main Effects - Does the decision aid of providing shot opportunity information have an impact, regardless of the quality of the air picture? And, does automated weapon target assignment have an impact, regardless of the quality of the air picture? 3) Automated Aiding X Air Picture Interaction - Does the impact of either type of automated aiding only emerge as a function of the level of air picture quality? Is either type of automated aiding more susceptible to degradation of the air picture than the other?

Direct Assessment Methods
Structured Operator Interviews

Test Results with Direct Assessment by Warfighters

- **Operators Rate Compensation Required to Execute Each Cell in the Design Matrix**
    - **Modified Cooper-Harper Rating Scale**
    - **Administered After All ▮ Air Picture Quality Levels Run for ▬▬ Case**

- **Diagnostic Explanations of:**
    - **Types of Compensatory Actions**
    - **Situation Awareness**
    - **Decision Confidence**
    - **Effectiveness / Adherence to Firing Doctrine**

Examine Impacts of Independent Manipulations (main effects & interactions) by:
- **C2 Level/Position:**
    - **Area Air Defense Commander (AADC)**
    - **Regional Air Defense Commanders (RADCs)**
    - **Tactical Action Officers (TAOs)**
    - **Tactical Directors (TDs)**
- **Individual Warfighter**
- **Centralization X Air Picture Quality Effects**
- **Automation X Air Picture Quality Effects**

Text

Direct Assessment Methods
Structured Operator Interviews

Method: There were five interviews conducted with each of the key players; one interview for each of the five cases. Each interview was done after all three air picture quality levels had been run for a particular case. An interview consisted of the operator rating the amount of compensation (see following rating scale description) for each of the three air picture quality levels. The operators were then asked what kinds of compensations were required (for each air picture level) to execute the mission, given the doctrinal and other conditions associated with the case. Operators were then asked what impact there was on mission effectiveness.

Data: Though subjective, the data produced by the structured interviews were both quantitative and qualitative. It provided numerical and verbal input for each cell of the design matrix, and thus for each cell in each of the overlapping factorial designs. As such, the interview data allowed for a non-statistical examination of the main effects and interactions of the experimental factors varied in this study.

Rating Scale Definitions

1-3: "Air Picture Satisfactory Without Improvement"
- I have to do little or nothing for which to compensate

4-6: "Adequate Awareness & Confidence Attainable With a ███████ Amount of Compensation"
"Deficiencies Warrant Improvement"
- current systems against current threats is usually rated "4" or "5"
- I'm working; I'm winning.

7-9: "███████ Amount of Compensation"
"Deficiencies Require Improvement"
- risk goes up; confidence goes down
- ██████████████████
- I'm working beyond my limit; I'm losing.

10: "No Situation Awareness Is Possible"
"Improvement Mandatory"
"Decision Making is Impossible for Extended Periods of the Tactical Scenario"
**- ████████**

Text

Rating Scale

The rating scale is a modification of the Cooper-Harper rating scale, first developed by NASA for use by test pilots in evaluating flight control systems. The scale generally produces stable ratings because it involves a two-stage decision tree. The users answer high-level questions first, which channel them into small clusters of rating boxes. Within each cluster, the individual ratings have clear definitions on which to base a selection. Users are not just picking a number from one to ten. Another feature of the scale is that it assumes that the user is an expert in his or her domain, and can perform the task or mission with almost any equipment. The central question is how much the user has to compensate for the shortcomings of the equipment.

In this study, the term "compensation" refers to anything the operator has to do because of shortcomings in the air picture. Compensation may take the form of increased workload (cognitive, verbal), with increases in a variety of activities, e.g., voice comm, mental calculations. It may also include incurring risks or making decisions despite an inadequate situational awareness.

Area Air Defense Commander (AADC)
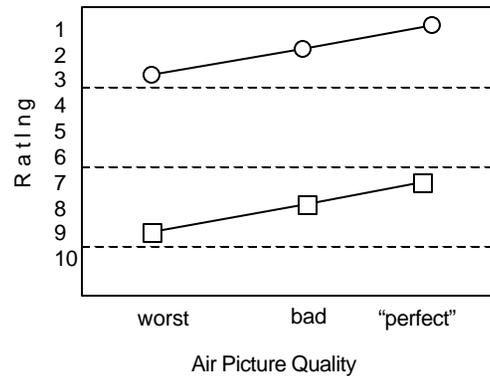
- **Very small but consistent main effect for Air Picture Quality.**

- **Zero interaction between Air Picture Quality and Shot Opportunity Info.**

- **Zero Interaction between Air Picture Quality and Degree of C2 Centralization.**

- **Huge main effect for Shot Opportunity Information.**

Text

Area Air Defense Commander (AADC)

██████████████, ██████████████████████, ██████████, ██
████████████: The AADC gave the same ratings to all four cases.
  - All ratings were very high. There was minimal to zero compensation
  - Voice communication was needed to tell organic tracks from false, and to tell who was engaging on organic tracks.

██: Not having the shot opportunity information had a huge impact. There was a minor impact of air picture quality.
  - Without it, he could not monitor employment.
  - He felt that he could only execute the mission because of trust in the three RADCs. That trust was built during prior training runs and during an earlier series of wargames.

C2 Centralization and Air Picture Quality Impacts
Regional Air Defense Commanders (RADCs)

Brown RADC

Rating (1–10), Air Picture Quality (worst, bad, "perfect")

Green RADC

Rating (1–10), Air Picture Quality (worst, bad, "perfect")

Blue RADC

Rating (1–10), Air Picture Quality (worst, bad, "perfect")
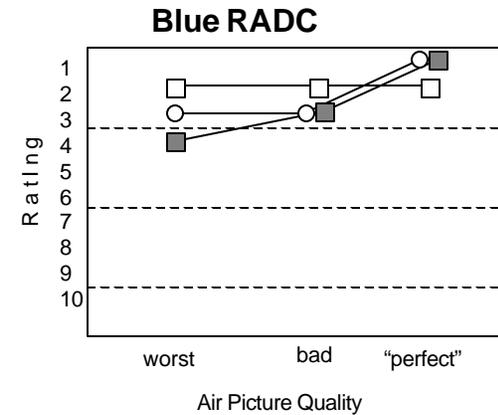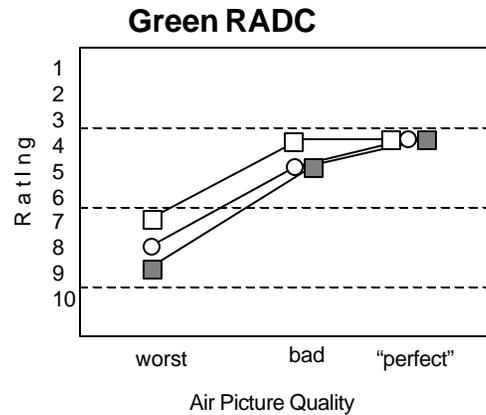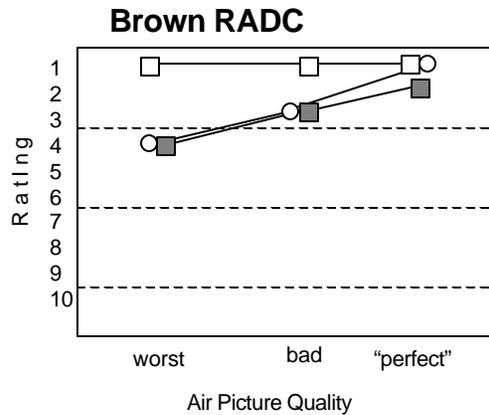
Key: Increasing Centralization

□ = Physical Deconfliction

○ = Decentralized w/ Upper Tier Coord.

■ = Centralized Control at RADC

- **Big Individual Differences**
- **Green RADC Dominated by "Pictorial" Needs**
- **Brown & Blue: Flexible, "Data" Types**

Text

C2 Centralization and Air Picture Quality Impacts Regional Air Defense Commanders (RADCs)

Brown & Blue: Very similar patterns of responses.
- Very small main effect for air picture quality levels.
- Slight Centralization X Air Picture Quality interaction

█████████████ - There was almost no impact of air picture quality. Ratings were consistently high. Compensation was not a factor.

███████████████████ - Low impact of air picture quality. Voice communication was needed to clarify shots, organic tracks, engagements.

██████████ - Low impact of air picture quality. "Comm intensive but organized."

Green: Very large impact of air picture quality levels in all three cases.
- Very large main effect for Air Picture Quality
- Slight Centralization X Air Picture Quality interaction

█████████████ - Degraded air picture quality required increased screen manipulation; resulting in an intolerable amount of compensation.

██████████████████ - Degraded air picture quality required increased screen manipulation. Having the shot opps window actually lowered situation awareness. Compensation was even more intolerable.

██████████ - The "shot opp list was a detractor". At the worst level of air picture quality, he had to pass authority down to the shooters to sort drops, etc..   Very high workload was required to have any situation awareness.

Automated Battle Management Aids & Air Picture Quality Impact
Regional Air Defense Commanders

Brown RADC

Green RADC

Blue RADC

(R a t i n g axis, values 1–10 for each chart)

Air Picture Quality — worst, bad, "perfect" (for each chart)
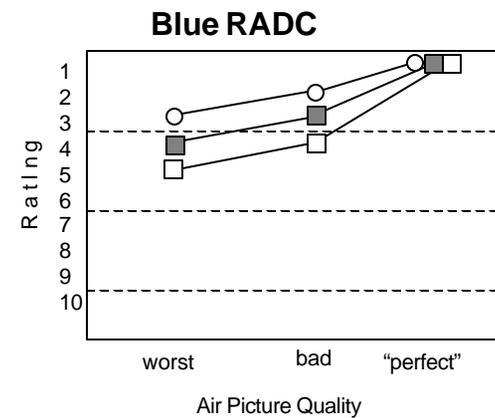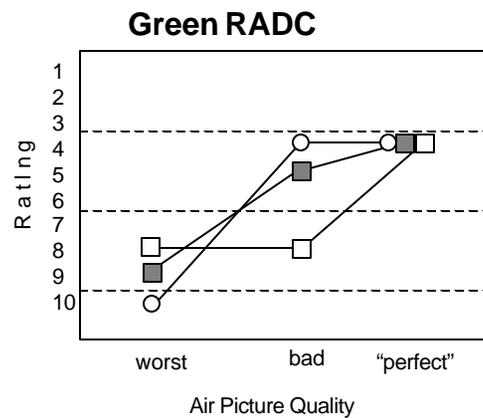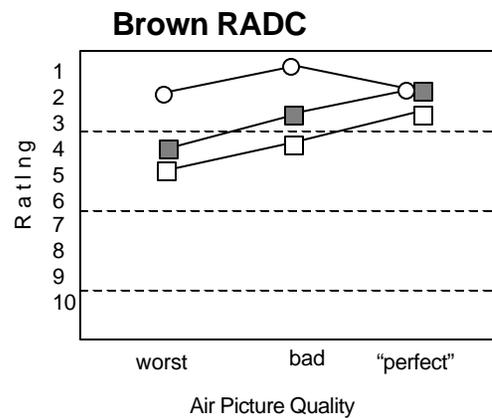
Key: All Centralized

☐ = Human, No Shot Opportunity Info.

■ = Human, With Shot Opportunity Info.

○ = Automated Battle Manager doing
Weapon Target Assignment, with
Human Override Possible

- **Big Individual Differences**
- **Green RADC Pictorial Dominance Persists**
- **Green RADC Need to Control Automation**
- **Degraded Air Picture X Automation Interaction**

Text

Automated Battle Management Aids & Air Picture Quality Impact
Regional Air Defense Commanders

Brown & Blue: Very similar patterns of responses.
• Very small main effect for Air Picture Quality levels: ratings drop from near-perfect to tolerable
• Slight Automation X Air Picture Quality interaction, but no cross-overs: ABM best, no Shot Opps Info worst at the worst level of air picture quality.
██████████████ - Low impact of Air Picture Quality levels. Ratings were in the highly tolerable range. The need for voice comm went up; and they had to push down more initiative to the shooters.
        - "[shot opps info] not a big deal"; not having it declutters the scope
██████████████ - Low impact of Air Picture Quality levels. Voice communication was intensive but organized.
█████████ - Low/zero impact of Air Picture Quality levels. Interface workload was lower. Compensation was minimal. Voice comm was needed to check organic status of tracks.

Green: Very large impact of Air Picture Quality levels in all three cases.
• Very large main effect for Air Picture Quality: Ratings dropped from high tolerable to intolerable/total chaos.
• Big ABMA X Air Picture Quality interactions: ABM worst, No Shot Opps Info best at" worst" air picture level; reverse order at "bad" level.
██████████████ - big impact from perfect Air Picture to "bad" level; no change at worst level. He had to push down judgments/sorting to shooters. He maintained effectiveness; but, lost efficiency; and lost advantage of centralized battle management.
██████████████ - He found the "shot opp list a detractor", which pushed him past his "break point" when the air picture was the most degraded. He had to pass authority to shooters to sort drops, etc. He had very high workload to achieve any situation awareness.
█████████ - No change in ratings from perfect air picture to"bad" level; but, there was a huge increase in rated compensation at the "worst" level. He was trying to "monitor the ABM and battle picture"; and had to "focus on the shot opp list instead of the big picture".

Inter-Operator Variability: Uniformity of Execution
Regional Air Defense Commanders

Key:

☐ = Physical Deconfliction

◯ = Decentralized w/ UT Coord.

▣ = Centralized Control

Centralization X Air Picture

Range

9 8 7 6 5 4 3 2 1 0

worst · bad · "perfect"

Air Picture Quality

Automation X Air Picture

Range

9 8 7 6 5 4 3 2 1 0

worst · bad · "perfect"

Air Picture Quality

Key:

☐ = Human Control, No Shot Opps Data

▣ = Human Control, With Shot Opps Data

◯ = Automated Battle Manager for WTA
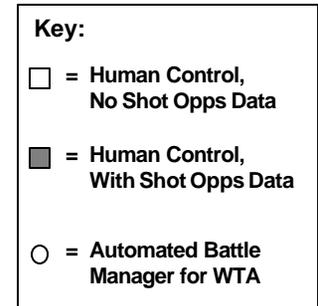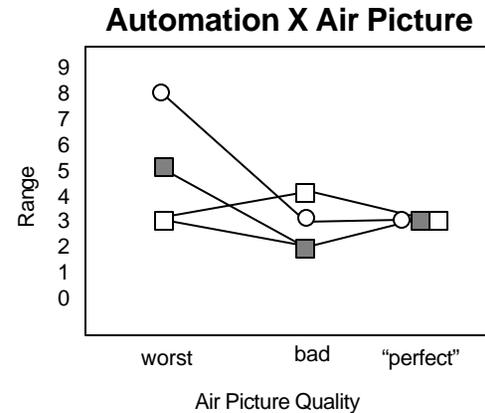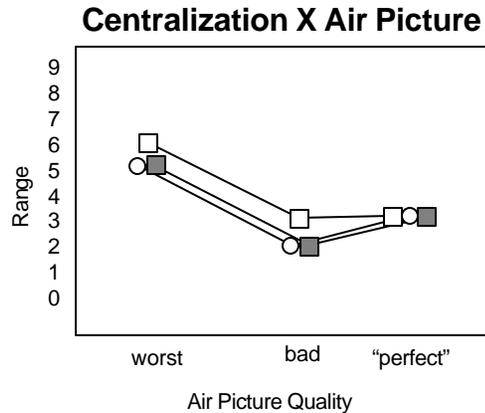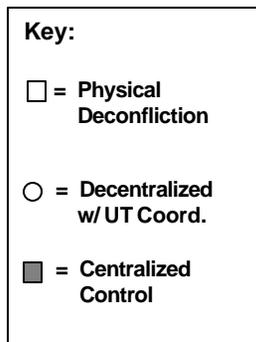
Plots show ranges of ratings differences among the three RADCs

- Degree of Centralized Control does not Cause Operator Differences

- Only Air Picture Quality Separates Players

- Automation is the Biggest Driver of Inter-Operator Variability

- . . . But, Only When Uncertainty Is High

Unambiguous Knees in the Curves

Text

Inter-Operator Variability: Uniformity of Execution
Regional Air Defense Commanders

Centralization X Air Picture Quality: The patterns and levels of inter-operator variability were virtually identical across the three different levels of C2 centralization. The only factor which contributed to greatly increased differences among the operators was air picture quality. Though fairly low, at the "perfect" and "bad" levels, inter-operator variability became extremely high when the air picture was at the highest level of degradation. This effect is attributable to the Green RADC's dependence on the air picture and his trying to clean it up. This pictorial dependence was the same whether he had to make weapon assignment (centralized control) or not (decentralized with upper tier coordination; physical deconfliction). These results show a very strong main effect of air picture quality when human operators are in control, regardless of how centralized doctrine is.

Automated Battle Management Aiding X Air Picture Quality: When the air picture quality was "worst", increasing the amount of automated aiding (from none to fully automated weapon target assignment) systematically increased the level of inter-operator variability. With no aiding (no shot opps info), differences among the operators was low and constant across the three levels of air picture quality. Having the shot opps info (middle level aiding) increased inter-operator variability at the "worst" air picture level. Full automation (ABM/WTA) resulted in the greatest level of inter-operator variability in the most degraded air picture condition. These results show a rather prominent interaction between air picture quality and degree of automated battle management aiding.

Tactical Action Officers (TAOs) & Tactical Directors (TDs)

Centralization X Air Picture Quality Effects

- **Low to Moderate Impact of Air Picture**

- **Minor Individual Differences**
  - **one TAO more susceptible to air picture degradation when upper tier coordination was required**

Automated Aiding X Air Picture Effects: Much Larger Individual Differences

- **TAO-2 similar to Green RADC**
  - **need to "control" automated battle manager**
  - **sensitive to air picture degradation**
  - **big, negative interaction between the two factors**

- **TAO-3 more sensitive to "no shot opps" when air picture was "worst"**

Text

Tactical Action Officers (TAOs) & Tactical Directors (TDs)

Centralization X Air Picture Quality Effects: At this level in the C2 structure the differences among individuals was not as striking as at the RADC level. Also, the impact of air picture degradation was low to moderate. The only rating which went into the "intolerable" range was given by one of the TAOs when upper tier engagement coordination was required (the duty of the TAOs and TDs) and the air picture was "worst".

Automated Battle Management Aiding X Air Picture Quality: In these conditions, one of the TAOs was considerably more susceptible to air picture degradation than the other operators at this level. He was also concerned with trying to control the automated battle manager. The combined effect of these factors resulted in a rating of "9" (just above "total chaos") when the automation was running and the air picture was "worst". In these regards, this TAO's response was similar to the Green RADC. The other TAO seemed more concerned with data, and only gave an "intolerable" rating when the air picture was "worst" and the shot opportunity information was not available. The ratings given by the TDs were generally better. Only one "intolerable" rating was given: "worst" air picture; no shot opportunity information available.

Summary

Big Individual Differences in the Impacts of the Factors:
- Air/Space Track Picture Quality
- Trust in Automation & the Need for Control

Training Issues:
- Understanding Automation
- Coping Strategies, e.g.:
  - studying weapons' and sensors' ranges before having the shot opportunity information taken away

Future Convergent Methods & Data
- Subjective: Quantitative & Qualitative Data
- Empirical/Objective Quantitative Data
  - C2 Decision Latency
  - C2 Decision Accuracy

Text

Summary

Individual Differences: Trained, experienced warfighters differ greatly in their reaction to system performance factors. Some are more susceptible to degradation of the air picture, while others can rely on other means of assessing the tactical situation. Some learn to trust automation and use it to lower their workload, while others feel the need to try to control it. In two cases, warfighters who were dependent on the quality of the picture also tried to manage automation. These two factors interacted strongly in these individuals, resulting in greatly increased workload and very low situation awareness, when the air picture was degraded and the automated battle manager was running. The definition of system performance requirements needs to be sensitive to the very real differences among warfighters and how those differences can impact the uniformity with which a mission is executed.

Training Issues: Two training issues which emerged from this study are: 1) Understanding Automation. Warfighters need to know what the black box is doing and practice operating with and without automated assistance. 2) Coping Strategies. Several individuals developed ways of working around conditions of system degradation, before they happened. One example was a RADC who studied the locations and ranges of the sensors and weapons under his command. He anticipated the uncertainty of a degraded air picture and the absence of shot opportunity data, and was able to command effectively within a tolerable band of workload. Some individuals may be more pictorially dependent or feel the need for control, but effective training should be able to mitigate the negative impact of uncertainty and automation on these individuals.

Future Convergent Methods & Data: The human effects data collected in this study were entirely based on self assessments made during post-run interviews. The quantitative and qualitative data revealed some important effects and gave some explanations for those effects. Future studies should continue to use subjective assessment methods. But, the subjective data should be complemented by more objective, empirical measures of the impact of experimental factors on the human warfighters. Two empirical measures which will be automatically derived in future wargames are: 1) C2 Decision Latency - How long foes it take a commander to assign a weapon to a target? And 2) C2 Decision Accuracy - Was the correct weapon assigned to the target?