Modeling Performance in C4ISR Sustained Operations: A Multi-level Approach

Linda R. Elliott
Veridian Engineering
2504 Gillingham Drive, Suite 201
Brooks City-base, San Antonio, TX 78235
Linda.Elliott@brooks.af.mil

Michael D. Coovert
Department of Psychology, PCD 4118
University of South Florida
4202 E. Fowler Ave.
Tampa, Florida 33620-8200
coovert@luna.cas.usf.edu
813-974-0482

Christopher Barnes, 1st Lt, USAF
Air Force Research Laboratory
Warfighter Fatigue Countermeasures
Brooks City-base, San Antonio, TX 78235
Christopher.barnes@brooks.af.mil

James C. Miller
Air Force Research Laboratory
Warfighter Fatigue Countermeasures Division
Brooks City-base, San Antonio, TX 78235
Jcmiller@brooks.af.mil

Abstract

In this paper, we report methodology and preliminary findings focused on the application of multi-level modeling techniques to distinguish effects of sleep loss and task demands on individual and team C4ISR decision making, coordination, and performance over time. We focus our efforts on measurement and modeling. First, we describe aspects of C4ISR scenario development, to ensure (a) psychological fidelity and operational relevance, (b) elicitation and assessment of performance constructs of interest, and (b) equivalence in scenario task demands and difficulty. Sustained operations research is challenged by the need for repeated-measures assessment, while minimizing effects of practice or experience. Second, we describe aspects of cognitive performance based on a standard cognitive test battery. Third, we describe other assessments (e.g. NEO PI personality assessment, mood-state inventory, Stanford Sleepiness Scale, physiological indices) that will be included in an overall approach to modeling fatigue effects, using multi-level hierarchical modeling analyses.

Introduction

USAF command and control (C2) warfighters face increasingly complex environments that represent the essence of naturalistic decision making-- multiple demands for enhanced vigilance, rapid situation assessment, and coordinated adaptive response. There are many perspectives on naturalistic decision making, however, all would agree that naturalistic decision making contexts are typified by expert, complex, interdependent and dynamic decision making, often under conditions of time pressure and/or uncertainty (Beach & Lipshitz, 1993). Cohen, 1993; Klein, 1993; Mitchell & Beach, 1990; Orasanu & Salas, 1991; Orasunu & Connolly; 1993; Rasmussen,1993).

In tactical C2 situations, the focus is on dynamic battle management and time-critical targeting. Information updates may be from air or from ground sources. Coordination demand is high-- reconnaissance and resource allocation depend upon close coordination between ground and air forces in a distributed network system of systems. Situations requiring close coordination and adaptive replanning are increasingly prevalent and challenging. Examples from the current situation in Iraq include tactics such as battlefield interdiction and close air support in situations requiring rapid movement of troops and armament. It is clear that challenges within these battle scenarios are considered critically important to air and ground superiority.

Much effort has been focused on the development of advanced technology to provide and represent time-critical information during mission execution. These capabilities are needed to facilitate, even enable, situation awareness and coordinated response in conditions of information complexity and time pressure. However, technology can only support, not replace, the role of the war fighters. Advanced technology enables closer coordination and accuracy of long-range weapons—yet it also increases the demand for human performance--for close team and multi-team coordination, shared situation awareness across numerous and diverse units, and rapid replanning across units, in hostile, dynamic, time-critical, and long-duration situations. In fact, technology increases the role and demands of the human decision maker.

Soldiers and leaders make decisions and responses in time-critical situations. While advanced technology affords paradigm shifts in information technology, it cannot replace C2 decision makers or troops on the ground, who must make tactical decisions under duress, often for long periods of time. Despite any particular advanced technology, individual performance will still vary, depending on the competence of each individual with regard to situation demands. We need to determine how to enhance the processes by which war fighters recognize, interpret and respond effectively in these situations.

Soldiers and leaders make decisions and responses in team situations. In these situations, effective response inevitably involves interaction with others: typically, information and/or resources must be distributed effectively and actions/events must be sequenced. Current military scenarios comprise operators and technology in a complex, dynamic, and interdependent system. Complex coordination must include adaptive problem solving. Prime examples include dynamic re-allocation of assets, for purposes such as retargeting and search-and-rescue.

Soldiers and leaders make decisions and responses in sustained situations. Sustained operations are integral to command and control—combat missions require vigilance over time and adaptive performance under stress. During the early stages of actual scenarios, members of the command center are often up for several days with little if any time for recuperative sleep. Over time, chronic fatigue will affect everyone, and the likelihood of error will increase. This is particularly relevant to C2 situations, which require constant monitoring, even when events are still.

Sustained operations result in acute and chronic fatigue effects. There is extensive documentation on the negative impact of acute and/or chronic sleep loss. In a review of findings, Bonnett (2000) report an array of negative effects. These effects include mood changes, disorientation, irritability, perceptual distortions, hallucinations, difficulty in concentration, and/or paranoid thinking, depending on the extent of sleep loss. Negative effects have also been demonstrated on a range of cognitive tests, such as monitoring tasks, speed / accuracy tests, short-term memory, logical reasoning, and mental subtraction/addition. Physiological effects are reflected in a variety of physiological tests, such as EEG, nystagmus, hand tremor, slurring of speech, sluggish corneal reflexes, hyperactive gag reflex, and increased sensitivity to pain.

While extensive data are available on effects of sleep loss on physiological, attitudinal, and cognitive function (Kryger, Roth, & Demnet, 2000), very few studies reported data regarding sleep loss effects on particular aspects of information processing in complex decision making tasks (Mahan, 1992; 1994).. Even fewer have reported on effects on team performance; however, a few prelimary studies, based on team simulation-based performance, provide some introductory results (Mahan, et al.,1998; Elliott et al., 2003). To continue this stream of research, the Chronobiology and Sleep Laboratory at Brooks City-base, San Antonio, TX has initiated a program of research on effects of sleep loss on information processing, communication, coordination, and decision making in complex simulation-based tasks. In this paper we shall focus on fundamental issues related to (a) elicitation of performance, (b) measurement, and (c) multi-level analyses that accommodate data complexity.

Multi-Level Modeling

Many kinds of data, including observational data collected in the human and biological sciences, have a hierarchical or clustered structure. For example, animal and human studies of inheritance deal with a natural hierarchy where offspring are grouped within families. Many designed experiments also create data hierarchies, for example clinical trials carried out in several randomly chosen centers or groups of individuals Multilevel models consider the fact of such hierarchies. We refer to a hierarchy as consisting of units grouped at different levels(Goldstein, Healy, & Rasbash, 1994).. In this case, we will have individuals clustered into teams, and we wish to model effects of fatigue over time, on various aspects of cognitive performance and operational decision making.

Before multilevel modeling became well developed as a research tool, the problems of ignoring hierarchical structures were reasonably well understood, but they were difficult to solve because powerful general purpose tools were unavailable. Special purpose software, for example for the analysis of genetic data, has been available longer but this was restricted to 'variance components' models and was not suitable for handling general linear models.

Multilevel modeling is particularly suited to fatigue research due to the necessity of repeated measures testing. Hierarchically structured data also occurs when the same individuals or units are measured on more than one occasion. A common example occurs in studies of animal and human growth. Here the occasions are clustered within individuals that represent the level 2 units with measurement occasions the level 1 units. There is a considerable past literature on procedures for the analysis of such repeated measurement data which has more or less successfully confronted the statistical problems. It has done so, however, by requiring that the data conform to a particular, balanced, structure. Broadly speaking these procedures require that the measurement occasions are the same for each individual. This may be possible to arrange, but often in practice individuals will be measured irregularly, some of them a great number of times and some perhaps only once. By considering such data as a general 2-level structure we can apply the standard set of multilevel modeling techniques that allow any pattern of measurements while providing statistically efficient parameter estimation. At the same time modeling a 2-level structure presents a simpler conceptual understanding of such data and leads to a number of interesting extensions. It also enables analyses of datasets with missing data, without “throwing out” any of the data.

Theoretical Approach

Figure A provides a representation of our overall approach to constructs, measures, and relationships, across a sequence of studies.

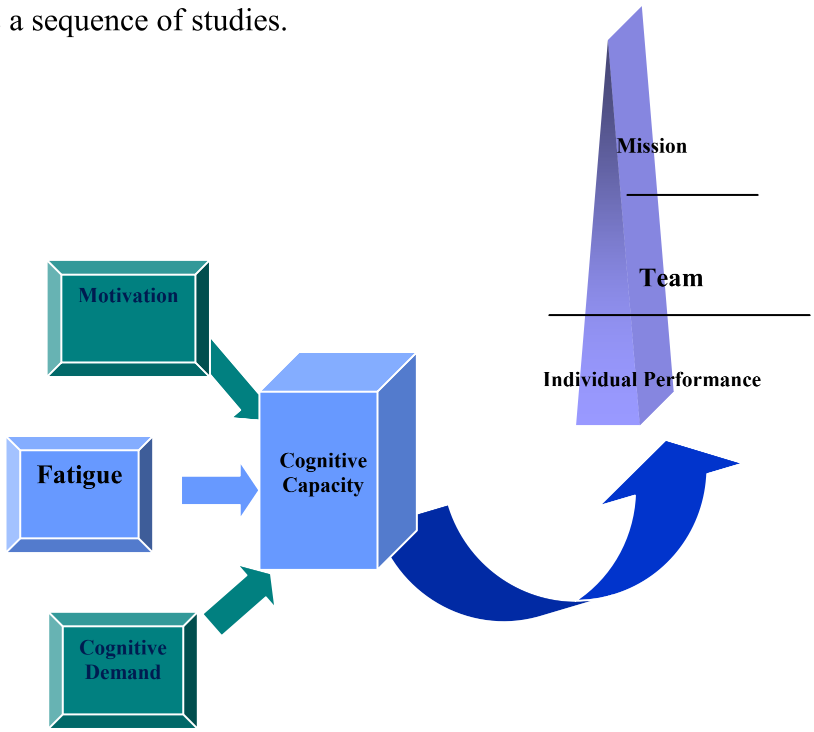


Figure A. Overall model of relationships among predictor and criterion variables of performance.

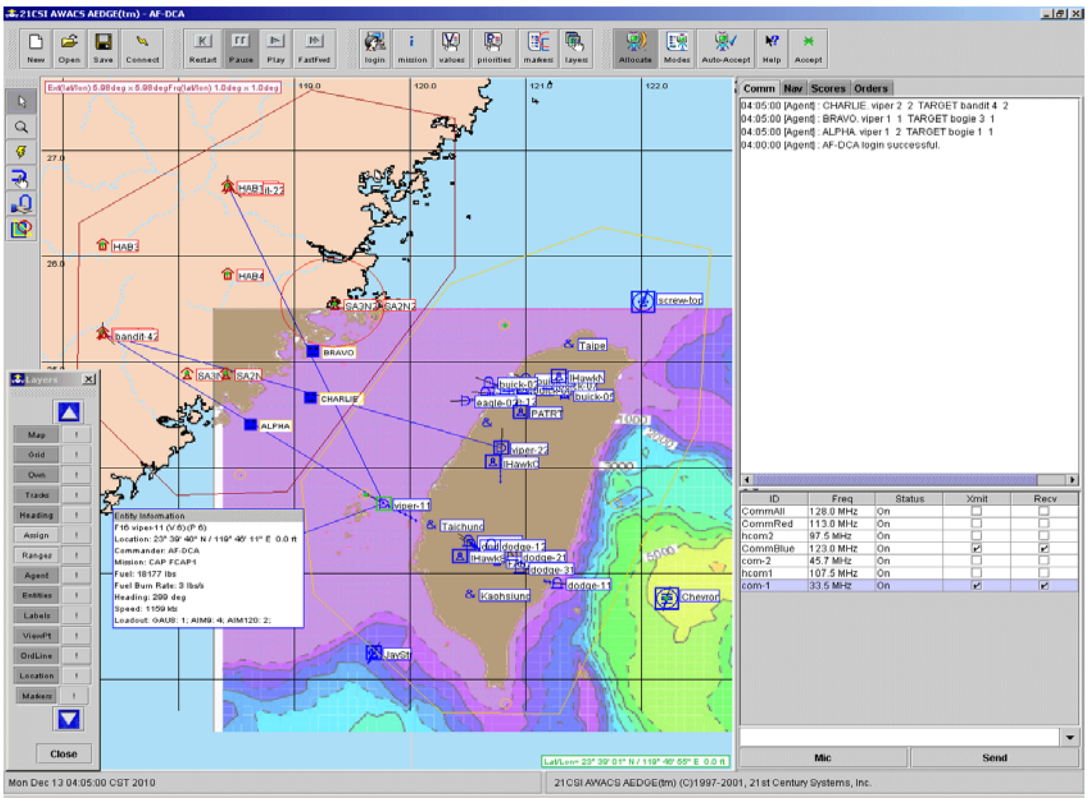
Simply put, the model predicts that fatigue will interact with cognitive demand to influence decision making and mission performance. More specifically, cognitive demands are expected to utilize cognitive resources from individual cognitive capacity (knowledge and ability), consistent with resource allocation models such as the Kanfer-Ackerman model of learning and motivation (Kanfer & Ackerman, 1989; Kanfer, 1991)). An underlying and general assumption is that fatigue is expected to reduce individual cognitive capacity. As this capacity is reduced, performance will be affected negatively, with regard to individual and teamwork performance. Motivation moderates the relationship between capacity and performance. For example, high capacity is not expected to result in high performance under conditions of low motivation. In contrast, high motivation is expected to ameliorate immediate effects of fatigue. Also consistent with literature on expert performance, it would be expected that experts would be more resistant to initial fatigue effects due to less effortful and more automatic, recognition-based cognitive processes.

In the overall model, fatigue diminishes total cognitive capacity, with increasing decrement over time. This systems view is consistent with quantitative research on effects of fatigue and chronobiology which supports the SAFTE (Sleep, Activity, Fatigue, and Task Effectiveness) model, which outlines effects of fatigue and chronobiology in more specific detail. (). (Eddy & Hursh SR,2000; Hursh, 1998)

Elicitation of Performance

Preliminary data collection will occur using a PC-based analogue of command and control functions, referred to as the AEDGE (Agent-Enabled Decision Group Environment). The AEDGE was developed based on cognitive and functional analysis of C3 mission, tactics, team member roles, and role interdependencies (Chaiken et al 2001, Barnes et al, 2002). Tactical scenarios were developed to capture core team coordination, decision-making and problem-solving task demands. Platforms such as the AEDGE provide an advanced PC-based platform for research and/or training, with high experimental control, manipulation, and online performance monitoring capabilities. The advantages of these capabilities are increased experimental control, manipulation, and operational relevance (Bowers, Salas, Prince, & Brannick, 1992; Cannon-Bowers, Burns, Salas, & Pruitt, 1998; Coovert, Craiger, & Cannon-Bowers, 1995; Schiflett & Elliott, 2000).

This study was very focused on transition of research to military operations. Investigations are based on an advanced agent-based synthetic team task environment (STE). STE platforms are PC-based platforms for training and performance research that are cognitive-based analogues of operational performance domains (Schiflett & Elliott, 2000). They provide realistic, complex and dynamic scenarios based upon the cognitive task demands of a particular operational domain, such as command and control (C2). Additional agent-based capabilities enhance scenario generation, manipulations, online assessments of performance, and data output formats. More extensive detail on the platform and software architecture are available elsewhere (Petrov et al,2000; Hicks et al. 2001; Kreft et al., 2000) . Figure 2 provides a representation of the interface.



Mission scenarios were typified by a strong demand for communication, shared awareness, coordinated action, and adaptive response to time-critical situations. Scenarios requiring dynamic replanning were carefully constructed to ensure equivalence in task demand and difficulty. This is particularly critical and challenging within this repeated-measures context. Two critical issues must be addressed: that of fidelity and equivalence of scenarios and event-based measures.

Fidelity. The fidelity of a system includes more than physical realism. The system must fulfill research or training goals through careful construction of scenarios tailored to meet specific needs—be they for theoretical research, performance modeling, and/or training. We have pointed out issues regarding multidimensional aspects of fidelity (Elliott et al, 2001; in press). Realistic equipment will certainly facilitate training and transfer of knowledge and procedural skill in equipment use. However, there are higher-level knowledge and skills that are independent from equipment procedures per se. These include knowledge of performance or training goals; individual roles and responsibilities; and perhaps multi-operator tasks and strategies. These aspects of expertise in a domain can be elicited, assessed, and/or trained without full equipment realism. *Indeed, the fidelity and validity of a fully realistic simulation system is also limited by the degree of realism and operational relevance of its scenarios.*

Previously, we have discussed 4 aspects of scenario fidelity (Elliott et al., in press, 2001). They are based on physical characteristics (i.e. degree of equipment realism), functional characteristics (i.e. capture of essential mission goals, roles, responsibilities, interdependencies, taskwork, and tactics), cognitive characteristics (e.g. cognitive demand characteristics such as difficulty, ambiguity), and the particular performance constructs that are elicited by the scenario (e.g. degree of coordination, task sequencing, adaptive problem solving required by the task).

For our scenarios, equipment realism is not high. Systems are PC-based and complex equipment procedures were simplified by menu-driven mouse clicks. However, functional fidelity is quite high. Scenarios were based on the functional roles and tasks of actual operations. The cognitive demand regarding the type and timing of tactical decisions is also quite high. Finally, the scenarios elicit the performance constructs of interest—that of individual and team decision making, coordination, and communication within tactical events.

Equivalence. Sustained operations research has particular demands with regard to repeated-measures. Measures have to be repeated over time in order to ascertain effects of fatigue. However, measures often cannot be replicated because of the need to minimize practice or learning effects. Even relatively simple cognitive tests that assess reaction time, working memory, or attention-switching require preliminary training to asymptote performance prior to the experimental session. Measures of more complex performance, such as logic or problem solving, are more difficult to assess over time, as most available tests do not have many equivalent forms. For many types of problems, repetition will elicit recognition-based performance: participants will be more likely to increase performance because they remember the problem.

Performance in the C4ISR scenarios will also increase, if we repeatedly use the same scenario. This prevents the assessment of fatigue effects. Once participants realize the same scenario is repeated, they will anticipate events and create strategies to increase performance while minimizing effort. In fact, in a recent study using these types of scenarios, 6 different scenarios were present to subjects over 5 experimental overnight conditions that were scheduled over a period of 6-8 weeks. While fatigue effects were clearly demonstrated in the first 2 experimental sessions, effects declined due to increased familiarity of scenarios and knowledge of strategies to perform more effectively with less effort. In fact, by the 5th experimental session, all subjects performed equally well regardless of who they were (individual differences were not significant) or time of session (Elliott, Coovert & Miller, 2003).

In the current study, each team of participants will experience only one overnight session. During the session they will perform in 8 different scenarios. The challenge inherent in this experimental design is the requirement of equivalence in scenario difficulty. We cannot confound results with scenarios varying in workload complexity or demand, and it can be quite difficult to craft scenarios that result in similar mean outcome scores.

Equivalent scenarios were constructed by assuring that all scenarios had (a) similar roles, (b) equivalent friendly assets, (c) equivalent hostile assets, (d) equivalent timing and tempo of events, (e) equivalent timing and tempo of additional hostile and friendly assets, and (f) equivalent geographic distances between hostile and friendly assets. Geographic distances affect the timing of hostile-friendly encounters and thus affects the tempo of workload demand. Appendix A has a listing of scenario assets for scenario 1 and 6, to illustrate the notion of equivalence in assets per role.

Each scenario had an ISR (Intelligence, Surveillance, and Reconnaissance), STRIKE, and SWEEP role, to be played by participants. Each role had similar assets and tactical goals. Assets were allocated across hostile and friendly roles in the same manner. For example, the ISR role would have the same number and type of Uninhabited Aerial Vehicle (UAV) assets at the beginning of each scenario, and would have additional assets appear at the same time through each scenario. He/she would face similar threat events, with regard to the number, type, and timing of hostile events. The same kinds of coordinating actions among the friendly roles are required in each scenario.

Recognition of the underlying "deep" structure of each scenario is minimized by changing the "surface" structure of each. One way this was achieved was by changing the geographic context and placement of assets. One scenario may be located in the geographic region of Taiwan, while another would be situation in Sri Lanka. The number and placement of assets would be equivalent, but not readily recognized. Another way this was achieved was by changing the type of hostile threat. In one version of the scenario, hostile threats were comprised of enemy surface-to-air missile sites. This situation is equivalent to a military tactic described as SEAD (suppression of enemy air defense). In another version, the hostile targets were theatre ballistic missile launchers. Identification and targeting of these targets are often referred to as "scud-hunting". The third version used in this study had hostile ships as enemy targets. Again, appendix A provides an example of how geography, targets, and asset labels change from scenario to scenario.

Scenario events were also timed to be equivalent. Assets appeared at particular times in each scenario. For example, in each scenario, hostile fighter aircraft appeared at specified times. The Appendix provides the chronology of events for scenario 1. Other scenarios have the same type and timing of events, where only the names of the assets change. Thus, in each scenario, the same cognitive and functional demands are presented to each role.

All scenario versions were created to have the same deep structure, in terms of workload demand and tactical principles. At the same time scenarios varied in surface characteristics to minimize recognition. In fact, experience in these scenarios would likely increase knowledge and application of general tactical principles that generalize over a variety of situations that appear to be dissimilar. For our study, learning was not something we wanted to elicit, therefore participants were trained as completely as possible. Participants experienced 34 hours of training before the experimental session.

Measurement of Performance

A variety of measures were collected, including individual physiology, self-report, performance on computer-based cognitive tests, and performance in the mission scenarios. In this paper we focus on assessments of team performance and their relation to criterion measures of performance.

Mission outcomes. Raw measures of mission outcome and team process were captured and time-stamped by the simulation. This includes descriptions and counts of events and actions, which then form the basis for various assessments of performance. For example, mission outcome scores were represented by the type, number, and relative value of assets that were lost, by "friendly" and "hostile" roles. Friendly assets included air bases, cities, surface-to-air missile launchers, uninhabited aerial vehicles, tanker aircraft, high-value reconnaissance aircraft, fighter aircraft, and bomber aircraft. Each asset was given a relative score value, generated by our weapons director expert, and validated by other experienced weapons directors. The loss of any friendly asset detracts from the score of the friendly team, and adds to the score of the enemy. In turn, hostile assets are similar. The loss of hostile assets adds to the score of the friendly team, and detracts from the score of the hostile. For these research participants, the overall mission outcome score was based on the point value obtained after subtracting all friendly "losses" from the total hostile "losses".

Observation-based measures provide additional data on individual and team interaction. Scenario observation sheets provide raters with descriptions of events as they evolve in the scenario. For example, in the first 6 minutes each scenario introduces several hostile threats. The observation sheets also describe what each team member should be doing at that time. During the first 6 minutes, the ISR role should be identifying decoys, to enable STRIKE to jam them and target them with bombers. At the same time, SWEEP fighters must protect STRIKE and ISR assets from air-to-air threat. Team members must act swiftly and in coordination with each other, or they will likely lose assets and waste resources.

Behaviorally-anchored rating scales based on descriptions of what each person should be doing are provided for time segments chosen based on the flow of events. Therefore, the appropriate and desirable responses to these events are also similar, across scenarios. For each introduction of assets and/or targets, the appropriate behaviors were identified for each role. Descriptions were generated for different levels of performance, leading to anchored scales for observer-based ratings of performance. Figure 3 provides an example of a tailored rating scale, this one tailored to rating the behavior of STRIKE, during the first six minutes of scenario 1.

0-6 min OVERALL rating of STRIKE (Circle)

During this time, Strike should be giving orders to bombers and jammers, jamming SAM sites, and assisting SWEEP and ISR.

1	2	3	4	5
POOR	FAIR	ADEQUATE	GOOD	EXCELLENT

POOR: no actions taken
FAIR: some action, very ineffective
ADEQUATE: some assigning of assets, no attention or assistance to others
GOOD: assets assigned, jamming accomplished, some coordination with others
EXCELLENT: assets assigned, jamming accomplished, assistance and coordination with others

Justify poor or excellent rating:

Figure 3. Observation-based rating scale for Strike, Scenario 1, 1st time period

Observation-based ratings of team communication. For specified time periods, based on events, observers provided ratings as shown in Figure 4:

0-6 min OVERALL rating of Team Communications (Circle)

1	2	3	4	5
POOR	FAIR	ADEQUATE	GOOD	EXCELLENT

POOR: No comm (or inaccurate comm) on conditions of friendlies/ no comm on coordination needs or activities
FAIR: Partial comm on conditions of friendlies / hostiles; some comm on simple coordination needs / activities
ADEQUATE: Some comm on conditions of friendlies / hostiles; some comm on ordinary coordination needs/activities
GOOD: Comm on conditions of friendlies/hostiles covered the majority of the situation; comm on more complex coordination needs/activities
EXCELLENT: Comm on conditions of friendlies covered the entire situation; comm on conditions of hostiles covered the entire situation; comm on the most complex coordination needs; comm to accomplish the most complex coordinations

Justify poor or excellent rating:

Figure 4. Rating scale for team communications

Audio capture of communications are another critical source of assessment. Communications were recorded in digital format to ease coding and analyses of data. Communications were initially coded for indications of teamwork, such as sharing of information or assets, sequencing actions, acknowledgements, requests for repeats, task-related encouragement, expressions of fatigue, and social comments (positive and negative). All comments were coded as to whether they requested or provided information.

Additional measures of individual characteristics include the Stanford Sleepiness Scale, the Profile of Mood States, the NEO-PI (all subscales), and performance on the ANAM cognitive test battery. The ANAM includes measures of reaction time, working memory, and multi-tasking ability. In addition, all subjects provided estimates prior to each scenario regarding the likelihood of attaining differing categories of performance outcomes, and afterward, their satisfaction with their outcomes.

Discussion

Data collection is still underway. Coding of communications have resulted in acceptable reliability (over 90% agreement), and refinements will likely increase agreement. Preliminary analyses on limited data have already revealed significant direct effects of fatigue on some criterion measures, even when power is very low (N = 4 teams, within-subjects time series analyses). Data from all sources will be ultimately analyzed using a multi-level hierarchical modeling approach that enables consideration of performance data of individuals, clustered in teams, to ascertain relationships and trends over time. A significant advantage to multilevel models is the ability to model and determine the significance of any cross-level interactions. A cross-level interaction is where a variable measured at one level (e.g. individual) interacts with a variable measured at another level (e.g., team). This is important as it allows us to assess and determine the significance of exactly the type of relationship of interest here. We will over time, focus on how individual performance affects team performance, and how both are affected by fatigue.

References

Barnes, C. M., Petrov, P. V., Elliott, L. R., & Stoyen, A. (2002). Agent based simulation and support of C3 decisionmaking: Issues and opportunities. Proceedings of the Conference on Computer Generated Forces and Behavior Representation, May 2002.

Beach, L. R. & Lipshitz, R. (1993). Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making. Norwood, NJ: Ablex Publishing Corporation.

Bonnet, M. H. (2000). Sleep Deprivation. In M. Kryger, T. Roth, & W. Dement (Eds.) Principles and Practices of Sleep Medicine, 53-68 Philadelphia, W.B. Saunders Company.

Bowers, C., Salas, E., Prince, C., & Brannick, M. (1992). Games teams play: A method for investigating team coordination and performance. Behavior Research Methods, Instruments & Computers, 24, 503-506.

Cannon-Bowers, J. A., Burns, J. J., Salas, E., & Pruitt, J. S. (1998). Advanced technology in scenario-based training. In J.A.Cannon-Bowers & E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp. 365-374). Washington DC: APA

Chaiken, S., Elliott, L., Dalrymple, M., & Schiflett, S., Coovert, M., Riddle, D., Gordon, T., Hoffman, K., Miles, D., King, T. (2001). Weapons Director intelligent agent-assist task: Procedure and findings for a validation study. Proceedings of the International Command and Control Research and Technology Symposium, June.

Cohen, M. S. (1993). The naturalistic basis of decision biases. . In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) Decision Making in Action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Coovert, M. D., Craiger, J. P., & Cannon-Bowers, J. A. (1995). Innovations in modeling and simulating team performance: Implications for decision making. In Guzzo & E. Salas (Eds.), Team Effectivesness and decisionmaking in organizations (pp. 149-203). San Francisco: Jossey Bass.

Eddy DR, Hursh SR (2000). Fatigue Avoidance Scheduling Tool. Phase 1 SBIR Final Report, NTI Inc., Dayton OH, to the Air Force Research Laboratory, contract F41624-99-C-6041.

Elliott, L. R., Coovert, M., & Miller, J. C. (2003). Ascertaining effects of sleep loss and experience on simulation—based performance. Proceedings of the 2003 Conference of the Society of Industrial/Organizational Psychology, Orlando, FL.

Elliott, L. R., Dalrymple, M. A., Schiflett, S. G., & Miller, J. C. (In Press). Scaling Scenarios: Development and Application to C4ISR Sustained Operations Research. In S. Schiflett, L. Elliott, E. Salas, & M. Coovert (Eds.) Scaled Worlds: Development, Validation, and Applications. Ashgate Publishing Limited, Surrey, England.

Elliott, L. R., Dalrymple, M., Regian, J. W., Schiflett, S. (2001). Scaling Scenarios for Synthetic Task Environments: Issues Related to Fidelity and Validity. Proceedings of the 2001 Meeting of the Human Factors and Ergonomics Society, September. Minneapolis, MN.

Goldstein, H., Healy, M. J. R., and Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, 13: 1643-1655.

Hicks, J., Stoyen, A., Zhu, Q. (2001). Intelligent agent-based software architecture for combat performance under overwhelming information inflow and uncertainty. Seventh IEEE International Conference on Engineering of Complex Computer Systems, 200-210.

Hursh, S.R. (1998) Modeling Sleep and Performance within the Integrated Unit Simulation System (IUSS). Final Report for the United States Army Soldier Systems Command; Natick Research, Development and Engineering Center, Natick, Massachusetts 01760-5020; Science and Technology Directorate; Technical Report: Natick/TR-98/026L.

Kanfer, R. & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative / aptitude approach to skill acquisition. Journal of Applied Psychology, 74, 657-690.

Kanfer, R. (1991). Motivation theory and industrial and organizational psychology. In M.Dunnette (Ed.), Handbook of I/O psychology (pp. 146-170).

Klein, G. A. (1993). A Recognition-Primed Decision (RPD) model of rapid decisionmaking. . In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) Decision Making in Action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Kreft, I, & De Leeuw, J. (2000). Introducing Multilevel Modeling. Thousand Oaks, Sage.

Kryger, M., Roth, T., & Dement, W. (2000). Principles and Practices of Sleep Medicine, 3rd Edition. Philadelphia: W. B. Saunders Company.

Mahan, R. P. (1992). Effects of task uncertainty and continuous performance on knowledge execution in complex decision making. International Journal of Computer Integrated Manufacturing, 5 (2), 58-67.

Mahan, R. P. (1994). Stress-Induces strategy shifts toward intuitive cognition: A cognitive continuum framework approach. Human Performance, 7, (2), 85-118.

Mahan, R. P., Elliott, L. R., Dunwoody, P., & Marino, C. (1998). The effects of sleep loss, continous performance, and feedback on hierarchical team decisionmaking. In (pp. 29-1-29-9).Edinburgh, Scotland.

Mitchell, T. R. & Beach, L. R. (1990). ...Do I love thee? Let me count ... Toward an understanding of automatic decision making. Organizational Behavior and Human Decision Processes, 417, 1-20.

Orasanu, J. & Salas, E. (1991). Team decision making in complex environments. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) Decision Making in Action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Orasunu, J. & Connolly (1993). The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) Decision Making in Action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Petrov, P., & Stoyen, A. (2000). An intelligent-agent based decision support system for a complex command and control application. Proceeding of the sixth IEEE international Conference on Engineering of Complex Computer Systems, Tokyo, Japan.

Rasmussen, J. (1993). Deciding and doing: Decision making in natural contexts. . In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) Decision Making in Action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Schiflett, S. G. & Elliott, L. R. (2000). Synthetic Team Training Environments for Command and Control. In Dee Andrews and Mike Mcneese (Eds.), Aircrew Training Methods. Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix A: Scenario Assets and Events.

Scenario 1. TAIWAN / Suppression of enemy air defense

(Boldprint indicates roles, assets and point values. Assets change in name, from scenario to scenario but not in function—see scenario 6).

MISSION (The only difference in mission among scenarios is the nature of the hostile targets, which were either surface-to-air missile sites, scud missile launchers, or hostile navy assets.

For scenario 1, the overall mission is to command and control the destruction of all Theater Ballistic Missile Launchers. Direct the assigned resources to find all the decoys, suppress the enemy air defenses, destroy all the missile launchers, and return all assigned assets to friendly territory.

The start of the scenario is on March 17th 2010 at 15:00 Zulu (Midnight Local Taiwan Time)

MISSION HVAA_Prime Protect ALL friendly resources against enemy attacks.
MISSION HVAA_Second Support Air Refueling Operations.

MISSION Strike_Prime Destroy enemy TBMs .
MISSION Strike_Second Use SEAD assets to protect friendly Sweep fighters with minimal losses.

MISSION Sweep_Prime Protect Strikers by sweeping hostile fighters from the sky.
MISSION Sweep_Second Use assets to help HVAA protect friendly bases and SAMs.

MISSION ISR_Prime Locate the mobile TBM threats.
MISSION ISR_Second Evaluate strike results.

GEOGRAPHIC LOCATIONS / CITIES (see Attached Map)
All scenarios shared the same number of friendly and hostile air bases, with equivalent distances among geographic sites.

Friendly and Hostile CAP Points

FCAP1	23°44’39”N119°37’16”E
FCAP2	25°20’29”N120°49’09”E
FCAP3	23°38’39”N121°47’03”E
FCAP4	21°45’50”N120°50’09”E
HCAP1	25°25’29”N119°42’15”E
HCAP2	23°55’27”N118°32’22”E

Friendly Airbases (1)

Kadena	26°10’27”N127°42’09”E
Sung	25°09’42”N121°44’51”E
Taoyuan	24°53’44”N121°00’44”E
Hsinchn	23°33’40”N120°21’59”E
Hwalien	22°34’57”N120°33’58”E
Taitung	22°50’44”N120°17’12”E
Pingtung	23°53’38”N121°22’06”E

Hostile Airbases

HAB1	26°39’39”N118°25’23”E
HAB2	25°16’37”N117°20’24”E
HAB3	26°06’32”N117°20’24”E
HAB4	25°49’44”N118°25’23”E
HAB5	28°19’29”N115°47’39”E
HAB6	25°16’37”N115°42’57”E
HAB7	22°46’52”N113°57’56”E
HAB8	27°29’34”N118°34’52”E

Cities	Allegiance	
Ma-kung	Hostile	24°45’32”N118°42’21”E
Taichung	Friendly	23°27’58”N120°10’49”E
Ciayi	Friendly	23°16’41”N120°17’12”E
Tainan	Friendly	22°54’09”N120°10’43”E
Pudtong	Friendly	23°40’39”N120°17’12”E
Taishsi	Friendly	24°00’37”N120°33’10”E
Kung-fu	Friendly	24°25’30”N120°39’34”E
Taipei	Friendly	25°10’14”N121°22’10”E
Kaohsiung	Friendly	22°48’36”N120°31’03”E

 # Hostile Missile Launchers (Point Value) All scenarios have the same number and timing of targets and target decoys.

DF-11a (25)	24°53’32”N118°24’23”E
DF-11b (25)	24°53’32”N118°34’22”E
DF-11c (25)	25°23’29”N118°17’24”E
DF-11e (25)	25°33’28”N119°20’18”E
DF-15b (50)	25°42’27”N119°02’19”E
(decoys)	
DF-11d (1)	24°48’32”N118°27’23”E
DF-11f (1)	24°48’32”N118°02’25”E
DF-11g (1)	25°26’28”N118°08’25”E
DF-15c (1)	24°38’33”N118°07’25”E
DF-15d (1)	24°38’33”N117°57’26”E

 # Hostile SAMs (associated with Missile Launchers, must be jammed/shot)

	Kill Radius(nm)	MaxHeight(ft)	MinHeight(ft)	Location
SA6-1 (1)	17	34,033	3,707	24°59’N118°26’E
SA6-2 (1)	17	34,033	3,707	25°38’N118°22’E
SA6-3 (1)	17	34,033	3,707	25°48’N118°47’E
SA-6-4 (1)	17	34,033	3,707	24°49’N117°47’E
SA4-1 (1)	16	18,840	304	24°49’N118°37’E
SA4-2 (1)	16	18,840	304	25°28’N119°22’E
SA-4-3 (1)	16	18,840	304	25°19’N119°09’E
SA-4-4 (1)	16	18,840	304	24°34’N118°12’E

FRIENDLY ROLES and ASSETS

HVAA—High Value Assets
MISSION HVAA_Prime Protect ALL friendly resources against enemy attacks.
MISSION HVAA_Second Support Air Refueling Operations.

SAM (Surface-to-Air) Sites

# Friendly SAMs	Kill Radius(nm)	MaxHeight(ft)	MinHeight(ft)
IHawkN (25)	22	25,500	300
IhawkC (25)	22	25,500	300
IhawkS (25)	22	25,500	300
PATRTN (50)	65	59,600	670
PATRTS (50)	65	59,600	670

TANKERS

CallSign	Type	Alt	Speed	Route	
Sunoco (100)	KC-10	29,000	450	25°28'N123°47'E	26°48'N125°02'E
Exxon (100)	KC-10	29,000	450	22°33'N122°04'E	24°33'N122°20'E
Texaco (100)	KC-10	29,000	450	22°14'N122°05'E	21°39'N119°27'E
CVN-72 (100)	Carrier	Surface	35	23°54'N122°05'E	24°39'N122°37'E

STRIKE Assets:
MISSION Strike_Prime Suppress enemy SAMs .
MISSION Strike_Second Use SEAD assets to protect friendly Sweep fighters with minimal losses.

	Type	ALT	Base	Speed	WEAPONS
Viper11 (15)	F18D	32,000	Kadena	620kts	M61A1 AIM9 2 AGM154 8
Viper12 (15)	F18D	32,000	Kadena	620kts	M61A1 AIM9 2 AGM154 8
Snake11 (15)	F18D	32,000	Kadena	620kts	M61A1 AIM9 2 AGM154 8
Snake12 (15)	F18D	32,000	Kadena	620kts	M61A1 AIM9 2 AGM154 8

	Type	ALT	Base	Speed	WEAPONS
Sting11 (15)	F18D	32,000	CVN-72	620kts	M61A1 AIM9 2 AGM154 8
Sting12 (15)	F18D	32,000	CVN-72	620kts	M61A1 AIM9 2 AGM154 8

	Type	ALT	Base	Speed	WEAPONS
Noisy51 (15)	Jammer	30,000	Kadena	29000	ALQ-99 1 ESM 1
Noisy52 (15)	Jammer	31,000	Kadena	29000	ALQ-99 1 ESM 1

	Type	ALT	Base	Speed	WEAPONS
Murky61 (15)	Jammer	31,000	CVN-76	580	ALQ-99 1 ESM 1
Murky62 (15)	Jammer	32,000	CVN-76	580	ALQ-99 1 ESM 1

Scenario 6 (CYPRUS, Sea Control)

Friendly and Hostile CAP Points

FCAP1	36°00’53”N031°45’49”E
FCAP2	34°49’00”N033°21’39”E
FCAP3	33°51’06”N031°39’50”E
FCAP4	34°48’00”N029°47’01”E
HCAP1	35°55’53”N033°26’39”E
HCAP2	37°05’47”N031°56’48"E

Friendly Airbases (1)		*Hostile Airbases*	
El-Sid	27°56’00”N034°11’38”E	HAB1	37°12’46”N034°40’50”E
Cretea	35°13’10”N033°50’49”E	HAB2	38°17’45”N033°17’48”E
Aetea	35°33’20”N034°24’45”E	HAB3	38°17’45”N034°07’43”E
Sparte	35°16’09”N033°04’41”E	HAB4	37°12’46”N033°50’56”E
Helles	35°04’11”N032°35’56"E	HAB5	39°50'30"N036°20'40"E
Kronos	35°05'58"N032°46'43"E	HAB6	39°55'12"N033°17'48"E
Gyrois	34°46'00"N033°24'39"E	HAB7	40°50'18"N030°48'03"E
Andersen	16°30'00"N017°32'00"E	HAB8	37°03'17"N035°30'45"E

Cities	*Allegiance*	
Troy	Hostile	36°55'48"N032°46'43"E
Hector	Friendly	35°02’23”N033°28’57”E
Paris	Friendly	34°46’00”N023°24’37”E
Kronosis	Friendly	35°10’58”N033°06’43”E
Inkarik	Friendly	35°20’57”N033°11’40”E
Ankara	Friendly	35°04’59”N033°31’39”E
Tuken	Friendly	34°58'35"N033°56'32"E
Cestiphon	Friendly	35°20'53"N033°31'22"E
Urbane	Friendly	34°57'07"N032°24'37"E

HOSTILE ASSETS / ASSIGNMENTS / INTEL

# Hostile Ships (Point Value)	Type	Location
Zeus08 (25)	Frigate	Tanoan Strait
Zeus01 (25)	Frigate	Tanoan Strait
Zeus02 (25)	Frigate	Tanoan Strait
Zeus03 (25)	Frigate	Tanoan Strait
Neptune02(50)	Destroyer	Tanoan Strait

(decoys)

Zeus04 (1)	Frigate	Tanoan Strait
Zeus05 (1)	Frigate	Tanoan Strait

Zeus06 (1)	Frigate	Tanoan Strait
Zeus07 (1)	Frigate	Tanoan Strait
Neptune01 (1)	Destroyer	Tanoan Strait

FRIENDLY ROLES and ASSETS

HVAA—High Value Assets

SAM (Surface-to-Air) Sites

# Friendly SAMs	Kill Radius(nm)	MaxHeight(ft)	MinHeight(ft)	Location
IHawkN (25)	22	25,500	300	35°39’N034°30’E
IhawkC (25)	22	25,500	300	35°15’N033°54'E
IhawkS (25)	22	25,500	300	34°52'N033°16'E
PATRTN (50)	65	59,600	670	31°33'N034°48'E
PATRTS (50)	65	59,600	670	34°54'N032°23'E

TANKERS

CallSign	Type	Alt	Speed	Route	
Petrol (100)	KC-10	29,000	450	33°33’N030°15’E	36°11’N029°40’E
Oiler (100)	KC-10	29,000	450	33°34’N030°34’E	33°18’N032°34’E
Pumper (100)	KC-10	29,000	450	31°54’N033°33’E	30°37’N034°49'E
CVN-66 (100)	Carrier	Surface	35	33°33’N031°55’E	33°01’N032°39’E

STRIKE Assets:

	Type	ALT	Base	Speed	WEAPONS
Dodge01 (15)	F18D	32,000	El-Sid	620kts	M61A1 AIM9 2 AGM-84 2
Buick01 (15)	F18D	32,000	El-Sid	620kts	M61A1 AIM9 2 AIM120 2
Dodge11 (15)	F18D	32,000	El-Sid	620kts	M61A1 AIM9 2 AGM-84 2
Buick11 (15)	F18D	32,000	El-Sid	620kts	M61A1 AIM9 2 AIM120 2

	Type	ALT	Base	Speed	WEAPONS
Hornet11 (15)	F18D	32,000	CVN-72	620kts	M61A1 AIM9 2 AGM-84 8
Hornet12 (15)	F18D	32,000	CVN-72	620kts	M61A1 AIM9 2 AGM-84 8

	Type	ALT	Base	Speed	WEAPONS
Clash11 (15)	Jammer	30,000	El-Sid	29000	ALQ-99 1 ESM 1
Clash12 (15)	Jammer	31,000	El-Sid	29000	ALQ-99 1 ESM 1
Gab11 (15)	Jammer	31,000	CVN-72	580	ALQ-99 1 ESM 1
Gab12 (15)	Jammer	32,000	CVN-72	580	ALQ-99 1 ESM 1

ISR Assets: UAVs & Net Players

CallSign	Type	Alt	Speed	Route

Sniffer	(100)	JSTARS	28,000	385	36°12’N029°48’E	34°04’N034°00’E
Sweeper	(100)	Rivet Joint	28,000	400	30°37’N034°49’E	34°06’N033°20'E

	Type	Alt	Speed	Location	FUEL(lbs)	WEAPONS
Foto11 (20)	UAV	10,000	200kts	27°16'N119°22'E	6000	CAM25 1
Foto 12 (20)	UAV	10,000	200kts	25°37'N119°42'E	6000	CAM25 1
Foto 21 (20)	UAV	10,000	200kts	23°57'N118°22'E	6000	CAM25 1
Foto 22 (20)	UAV	10,000	200kts	22°17'N118°22'E	6000	CAM25 1
Eyeball11 (60)	UAV	60,000	450kts	28°06'N119°05'E	9000	SAR 1
Eyeball12 (60)	UAV	60,000	450kts	20°47'N117°26'E	9000	SAR 1

SWEEP Assets: AWACS, F15s and F14s

CallSign	Type	Alt	Route
Blowout (100)	AWACS	29,000	33°25’N034°05’E35°00’N030°17'E

	Type	ALT	Base	Speed	WEAPONS
Tern11 (15)	F14	38,000	CVN-72	720kts	M61A1 AIM120 4 AIM7 2 AIM9 2
Tern12 (15)	F14	38,000	CVN-72	720kts	M61A1 AIM120 4 AIM7 2 AIM9 2
Raygun01 (15)	F15	40,000	El-Sid	720kts	M61A1 AIM120 4 AIM7 2 AIM9 2
Raygun02 (15)	F15	40,000	El-Sid	720kts	M61A1 AIM120 4 AIM7 2 AIM9 2

Appendix B. Content and timing of Scenario events.

Scenario 1.
0-6 minutes: tick1 - 36
Additional assets appear for Hostile (6 Migs) and Friendlies. Sweep gets 2 F15s, Strike gets 4 F-16 fighters and 2 jammers.

Min.Secs	
0.10	Hostile: 2 MIG31 (30) from HAB2**: Flanker11/12** to HCAP2 WEAPONS: FxFrdr 1 M61A1 1 AA12 4 AA11 2
0.20	Sweep: 2 F15 (15) from Kadena: **eagle21/22** ordered to FCAP2 WEAPONS :each has M61A1 AIM120 4 AIM7 2 AIM9 2
0.30	Hostile: Mobile Hostile SAMs(8) Pop-Up (see list of SAMs above)
0.50	STRIKE 4 F16(15) from Kadena, FCAP2: **Spy13/14**; Weapons: M61A1 AIM9 2 AIM120 2 ; 32,000ft; 620kts **; Zap13/14** Weapons: M61A1 AIM9 2 AGM88 4 ; 32,000ft; 620kts
1.00	Hostile 4 More MIG23 (15) **bandit-21/22**, HAB1 to HCAP1; **bandit-41/42**, from HAB2 to HCAP2
1.20	STRIKE Jammers EA-6Bs(15): from Kadena to FCAP2: **Jazz13/54**; 30-31,000ft; 580kts
6.0	Hostile: gain 4 MIG21 (10) (**Bandit81/82** from HAB1 to HCAP1; **Bandit91/92** from HAB2 to HCAP2)

6-15 minutes (36-90 ticks)

11.2 Hostiles gain 4 Mig23(15) (**bandit 51/52** from HAB1 to HCAP1; **bandit 61/62** from HAB2 to HCAP2) WEAPONS RADARx 1 M61A1 1 AA10 4 AA2 8

13.2 Hostiles gain 2 Mig31 (30) (**Flanker13/14** from HAB2 to HCAP1,2)
WEAPONS FxFrdr 1 M61A1 1 AA12 4 AA11 2

15.00 Sweep gain 2 F14 (15) **tomcat21/22** from CVN-72 to fCAP3)
WEAPONS M61A1 AIM120 4 AIM7 2 AIM9 2; 38.000ft; 720kts

15-25 minutes (tick 90-150)

17.0 STRIKE-- 2 F-18Ds, HARM Shooters (15) (**WASP33/34** fromCVN-72 to FCAP4)
WEAPONS M61A1 AIM9 2 AGM154 8; 32,000ft; 620kts

19.0 STRIKE-2 Jammers(15) (**Buzz13/14** from Kadena to FCAP2) WEAPONS ALQ-99 1 ESM 1; 30-31,000ft; 580kts

20.0 HOSTILE-4 MIG21 (10) (**bandit 83/84** from HAB1 to HCAP1; **bandit 93/94** from HAB2 to HCAP2) WEAPONS RADARx 1 M61A1 1 AA10 4 AA2 8

21.40 HOSTILE - 4 MIG23 (15)
(**Bandit7-1/7-2; Bandit 6-1; 6-2)**
WEAPONS < RADARx 1 M61A1 1 AA10 4 AA2 8

23.20 HOSTILE-4 MIG23 (15) (**Bandit3-1; 3-2; Bandits 5-1; 5-2)**
WEAPONS < RADARx 1 M61A1 1 AA10 4 AA2 8

24.10 HOSTILE- 2MIG31 (30)
(**Flanker15** from HAB2 to HCAP1; **Flanker16** from HAB2 to HCAP2)
WEAPONS FxFrdr 1 M61A1 1 AA12 4 AA11 2

25-40 minutes

25.00 HOSTILE—SAM SITES (all decoys, value = 1)

Hostile Missile Launchers (Point Value) (decoys)

SA2B (1) 24°48'32"N118°27'23"E
SA3B (1) 24°48'32"N118°02'25"E
SA6-1B (1) 25°26'28"N118°18'24"E
SA4-1B (1) 25°30'28"N119°22'17"E
SA10B (1) 25°37'27"N119°25'17"E

these should be identified by ISR
wasted assets if jammed or destroyed

26.0 If nuke isn't destroyed, it launches and kills Andersen AFB. WDs should return to original mission.

26.36 HOSTILE-- 5 SAM SITES (25,50)

\# Hostile Missile Launchers (Point Value)

SA2C (25)	25°23'29"N118°17'24"E
SA3C (25)	25°38'27"N119°37'16"E
SA6-1C (25)	25°16'29"N117°20'29"E
SA4-1C (25)	26°39'21"N118°25'23"E
SA10C (50)	25°36'27"N119°25'17"E

ISR Assets: UAVs & Net Players

MISSION ISR_Prime Locate the mobile TBM threats.

MISSION ISR_Second Evaluate strike results.

NET PARTICIPANTS

CallSign	**Type**	**Alt**	**Speed**	**Route**	
Sounder (100)	JSTARS	28,000	385	21°47'N119°27'E	25°59'N121°34'E
Cueball (100)	Rivet Joint	28,000	400	26°48'N125°02'E	18°44'N121°33'E

Sensor UAVs

	Type	Alt	Speed	Location	FUEL(lbs)	WEAPONS
Spot11 (20)	UAV	10,000	200kts	27°16'N119°22'E	6000	CAM25 1
Spot12 (20)	UAV	10,000	200kts	25°37'N119°42'E	6000	CAM25 1
Spot21 (20)	UAV	10,000	200kts	23°57'N118°22'E	6000	CAM25 1
Spot22 (20)	UAV	10,000	200kts	22°17'N118°22'E	6000	CAM25 1
Foto21 (60)	UAV	60,000	450kts	28°06'N119°05'E	9000	SAR 1
Foto22 (60)	UAV	60,000	450kts	20°47'N117°26'E	9000	SAR 1

SWEEP Assets: AWACS, F15s and F14s

MISSION Sweep_Prime Protect Strikers by sweeping hostile fighters from the sky.

MISSION Sweep_Second Use assets to help HVAA protect friendly bases and SAMs.

AWACS

CallSign	Type	Alt	Route
Checkmate (100)	AWACS	29,000	22°15'N120°38'E26°04'N122°13'E

	Type	ALT	Base	Speed	WEAPONS
Lynx41 (15)	**F14**	38,000	CVN-72	720kts	**M61A1 AIM120** 4 **AIM7** 2 **AIM9** 2
Lynx42 (15)	**F14**	38,000	CVN-72	720kts	**M61A1 AIM120** 4 **AIM7** 2 **AIM9** 2

	Type	ALT	Base	Speed	WEAPONS
Talon31 (15)	**F15**	40,000	Kadena	720kts	**M61A1 AIM120** 4 **AIM7** 2 **AIM9** 2
Talon32 (15)	**F15**	40,000	Kadena	720kts	**M61A1 AIM120** 4 **AIM7** 2 **AIM9** 2