16th ICCRTS: Collective C2 in Multinational Civil-Military Operations

Operational planning with uncertain and ambiguous information: Command and control and the natural environment

Topics: Topic 5: Collaboration, Shared Awareness, and Decision Making
Topic 6: Experimentation, Metrics and Analysis
Topic 7: Modeling and Simulation

Karl D. Pfeiffer, Susan G. Hutchins, David L. Kleinman, Scot Miller, and William G. Kemple
Naval Postgraduate School, Monterey, CA

POC: Karl D. Pfeiffer
Graduate School of Operational and Information Sciences
1411 Cunningham Rd.
Naval Postgraduate School
Monterey, CA 93943
831.656.3635
kdpfeiff@nps.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUN 2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE **Operational planning with uncertain and ambiguous information: Command and control and the natural environment**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School,Graduate School of Operational and Information Sciences,1411 Cunningham Rd,Monterey,CA,93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 16th International Command and Control Research and Technology Symposium (ICCRTS 2011), Qu?c City, Qu?c, Canada, June 21-23, 2011.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**44**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Operational planning with uncertain and ambiguous information: Command and control and the natural environment

Operational planners, particularly military planners, are often faced with constructing a plan using ambiguous data in highly complex or rapidly evolving situations. Environmental information represents a particular challenge for planners. The state of the art in geophysical fluid dynamics leaves significant uncertainty in forecast conditions. Even with perfect knowledge of the future state of ocean and atmosphere, translating these conditions into mission impacts can be difficult and result in ambiguity in interpretation. In this research, we examine the use of meteorology and oceanography (METOC) information by operational planners. An experiment was conducted using human subjects participating in a computer-mediated planning simulation. Player teams were charged with constructing plans to allocate assets to tasks in a five-day operational scenario. Players were required to integrate dynamic METOC information presented with varying levels of information richness (ambiguity in weather conditions) and varying levels of information structure (ambiguity in weather impacts). Plans were evaluated for both completeness and robustness, where robustness was assessed by considering the plan performance over the distribution of likely METOC conditions in the mission area. Results offer insight into the more effective employment of METOC personnel in the planning process, and into better presentation of METOC information to planners.

1. Overview

How do people make the best use of uncertain and highly perishable information? Operational planners, particularly military planners, are often faced with constructing a robust and effective plan using ambiguous data in a highly complex or rapidly evolving situation. Environmental information represents a particular challenge for planners. The state of the art in atmospheric and ocean sciences leaves significant uncertainty in forecast conditions. Even with perfect knowledge, translating these conditions into mission impacts can be difficult; delivering a more accurate weather forecast does not necessarily provide the planners with more useful information.

> *"Listen, S-2," the colonel said, "I don't care about how many inches of rainfall to expect. I don't care about the percentage of lunar illumination. I don't want lots of facts and figures. Number one, I don't have time, and number two, they don't do me any good. What I need is to know what it all means."*
> —USMC Doctrinal Publication 6 *Command and Control* (1996)

In the 2011 Model-Based Experiments on Adaptive and Scalable Architectures for Maritime Operations Centers experiment, we examine the use of meteorology and oceanography (METOC) information by operational planners, using a computer-mediated planning tool developed at the University of Connecticut (UCONN). The game scenario required players to consider METOC information while assigning assets to tasks over a multi-day period. Players constructed a series of plans based on the resources required to complete a task, the assets available, and the effectiveness of those assets under a range of predicted METOC conditions.

Information richness was manipulated as the independent variable across three levels ranging from low to high. Dependent measures examined the quality of the plans produced, the effective use of assets, and the time players spent in collaboration and deliberation over different aspects of the plan. Although an obvious hypothesis is that players given more and richer METOC information will produce better plans, our work sought to develop clear insight into both direction and *magnitude* of these differences.

The motivation for this work was, in part, that much of the current METOC support to military operations is accomplished with automated computer products and human weather specialists available remotely (for example, by text chat or telephone) to deployed forces; this is operationalized in the manipulation of information structure. While weather and ocean forecasts always carry some implicit or ambiguous uncertainty (for example, 60% chance of rain tomorrow), both the Navy and the Air Force are seeking mechanisms to evaluate explicitly the uncertainty attached to certain numerical weather prediction techniques. These clear confidence bounds on the forecast may improve the information given to military planners and decision-makers, though it is not clear that this will necessarily lead to better plans and decisions. This explicit uncertainty is operationalized in the manipulation of information richness, and experimental results help to clarify the connection between consideration of uncertainty and improved decision-making.

The remainder of this paper is organized as follows: Section 2 provides a brief background and review of the literature; Section 3 presents the experimental design and data collection plan; Section 4 discusses the qualitative and quantitative outcomes from the experiment; and Section 5 summarizes the results and conclusions, and then outlines avenues for future research.

2. Background

2.1 The Adaptive Architectures for Command and Control research program

The Model-Based Experiments in Adaptive and Scalable Architectures for Maritime Operations Center is connected to the larger Adaptive Architectures for Command and Control (A2C2) program, operating at Naval Postgraduate School for over 15 years. The A2C2 research program has developed a multi-disciplinary research agenda to conduct experimentation on issues critical to Maritime Operations Centers (MOC) and Maritime Headquarters (MHQ). Senior Navy leaders have stated a need to refocus and enhance the Navy's ability to function at the operational level of war (U.S. Fleet Forces Command, 2007). MOCs were conceived to enable these capabilities while providing a degree of standardization among the maritime headquarters (MHQ with MOC Concept of Operations, 2007). A MOC empirical research campaign is underway where the emphasis is on *operational versus tactical* activities, and *planning versus reacting*. Because of its complexity, its mission to oversee large operations, and its dynamic structure, the MOC is an ideal organization for research on organizational structure, C2, and the process of mission planning. Since the MOC was designed to effectively integrate the planning elements of Current and Future Operations (COPS and FOPS) to provide more rapid and accurate resource allocations that are consistent with mission requirements, our first experiment focused on the MOC with emphasis on intelligence, surveillance, and reconnaissance (ISR) (Hutchins et al., 2009).

A second experiment conducted in 2010 to investigate an issue relevant to MOCs. In this study, the MOC planning teams operated in either an integrated or isolated fashion. The integrated FOPS team was supported by a decision aid and planning tool that fostered coordination, while the isolated FOPS team used a planning tool with a reduced coordination capability. The objective of this experiment was to examine the potential problems that could arise when forming an operational planning team (OPT). An OPT is a task-organized team formed to conduct integrated planning for a specific mission. An OPT is often formed by the MOC because it offers the advantage of a focused group of subject matter experts approaching the problem in an integrated manner. However, problems may be associated with this team being somewhat isolated in situations that require the OPT to coordinate closely with the rest of the MOC. The overarching research question sought to understand how emergent events are best handled when resources must be shared among separate planning teams--for example when an OPT is formed.

The 2010 study was designed to examine the efficiency and planning performance of two alternative organizational structures: (1) *Integrated* – where planning teams plan with a real-time view of others' resource planning, and (2) *Isolated* – where planning teams operate in isolation, without the ability to directly view others' resource planning. The first experimental hypothesis was that "integrated teams create more effective plans than isolated teams due to their real-time awareness that enhances the interdependent solution. Our second experimental hypothesis was that isolated team members experience higher levels of workload than integrated team members because their lack of real-time planning status requires more frequent status-related communication in addition to collaborative effort. Our third experimental hypothesis was that isolated team members communicate more frequently in response to emergent events because isolated team members must communicate to learn how others alter plans in response to unexpected events.

Teams in the integrated condition produced higher quality plans than the isolated condition due to the enhanced shared situation awareness provided by the planning software and as expected, the average overall workload reported by isolated team members was significantly greater than workload reported by integrated team members across all four experimental sessions (Hutchins et al., 2010). The lack of shared situation awareness provided by their planning tool necessitated an increase in the amount of explicit (communication-based) status updates and coordination required to succeed on this interdependent task. Results suggest that as the scenario builds and the interdependence between task areas A and B became more complex, integrated team members perceived less mental workload and demonstrated more effective performance than isolated team members.

The objectives for the present study are twofold, to: (1) continue our model-based experimentation, and (2) explore new paradigms for empirical studies of critical issues appropriate for MOC laboratory research. Through the integration of analytical modeling, human-in-the-loop experimentation and computer simulation, our research has followed a "model-test-model-experiment" paradigm wherein models and associated simulations define and guide the experiments, and the results from the experiments are fed back to improve and enhance the models. In previous research we have focused on resource allocation in planning under different collaborative structures for sharing information. An implicit assumption in these studies was that this shared information was "perfect." In the present study we examine the use of imperfect, mission-critical information: characterization of the natural environment.

2.2 The ambiguity of forecasts for the atmosphere and oceans

Weather prediction is appealing as a purely deterministic, physical problem (Regnier, 2008b). This deterministic approach, however, is limited by physical realities in how well we can characterize the initial state of the atmosphere and oceans, and the degree to which we can carry this information through our time integration of the coupled differential equations that describe the atmosphere (Lorenz, 1963). This deterministic approach yields a single, physically plausible solution, though the accuracy of this solution begins to deteriorate quickly after about 144 hours in to the forecast under the best of conditions. Put simply, even with the technology and science we bring to bear in the 21st century, we are still challenged by accurate weather predictions beyond about six days.

Every numerical weather prediction carries an inherent uncertainty. Typically, the bounds of this uncertainty are unknown, though a skilled human weather forecaster can often interpret model output to attach a reasonable estimate of the uncertainty. Although this uncertainty estimate may be qualitative (e.g., the model often over-predicts rainfall amounts in the mountains), the human forecaster nonetheless can make this uncertainty clearer to decision-makers (Doswell, 2004).

Another approach to evaluating model uncertainty is to compare several models over the same forecast period, examining the mean and variance of the ensemble solution (Leutbecher & Palmer, 2008).

Regions of high variance in the ensemble may be interpreted as areas of significant uncertainty. The explicit numbers derivable from ensemble output provide another means to make clear to decision-makers how much trust to place in a particular forecast for a particular region. The interpretation of these numbers, however, is not always clear to decision-makers and may require some intervention by human weather forecasters (Morss, 2010).

2.3 Creating actionable intelligence from ambiguous predictions

No matter the skill with which we predict the future state of the atmosphere and oceans, these predictions are of themselves of little use to military planners. Decision-makers are largely concerned with when and to what degree their assets and capabilities will be affected by weather conditions (Moore et al., 2003). For planners with a trade-space spanning days or weeks, decisions to proceed with, accelerate or delay operations are connected to the expected weather conditions.

While at the tactical level it is common for operators to simply avoid or react to the weather, at the operational level planners may employ environmental forecasts to help ensure the best use of assets in time and space. Assuming that bad weather is transient, and particularly if the area of responsibility covers multiple weather regimes, planners can adjust assignments in concert with forecast environmental conditions. In bad weather areas, planners can often delay operations and reassign assets to operations in relatively good weather areas, increasing effort where the environment permits. These deferred tasks in initially bad weather areas may then receive increased effort when the weather improves. Effective planning in this case is connected to knowledge and *exploitation* of environmental forecasts.

2.4 Summary

One may argue that all planning takes place under uncertainty. The motivation for this work is that the natural environment and its inherent uncertainty represent both challenges and opportunities for the operational planner. The degree to which this information is useful to the planner, however, may depend upon the information content and structure. We next describe the research design for this investigation.

3. Method

3.1 Research question and working hypothesis

We seek to better understand how organizations employ perishable and uncertain information in the operational planning process. Does providing more information to planners lead to better planning? In the face of inherently uncertain information, if we make this uncertainty more clear through human intercession, or explicit quantitative bounds, will planners use and integrate this information in their deliberations?

This research is a natural evolution from previous A2C2 studies involving integrated planning teams in the US Navy Maritime Operations Center (Hutchins et al., 2009), most recently under different collaborative structures (Hutchins et al., 2010). These efforts were distinguished by computer-mediated experiments with human subjects playing the role of planning teams supporting the MOC. In the present study we examine the MOC planning process under a scenario where player actions are constrained by events in the natural environment (e.g., thunderstorms, heavy seas, dense cloud cover). Given the same environmental scenario, but markedly different information about the uncertainty in the natural environment, we ask: do teams integrate and apply this information effectively in their mission plans?

If we think of the experimental levels as ranging from low to high in terms of uncertainty information, we hypothesize that planning teams given high (richer) uncertainty information will outperform teams given low (less rich) uncertainty information. While this may seem obvious, our design is intended to address the perhaps more useful question: how much better do teams perform given richer uncertainty information?

This question is of significant operational relevance to both the Navy and Air Force, as there is a cost to keep humans deeply embedded in the forecast process, and a cost to produce explicit uncertainty bounds with numerical forecasts. Within the Department of Defense (DoD) the current trend is to consolidate METOC personnel in centers, typically located far from the forward edge of battle, and most often located in the continental United States (CONUS). Support to deployed operations is then provided with online product delivery and reach back service to these centers. Both the Navy (Sestak et al., 2008) and the Air Force (Nobis et al., 2008) are examining the use of ensemble numerical weather prediction to improve operational forecasts and improve the explicit uncertainty information attached to these forecasts. For both services, though, a lingering concern is whether forecasters and decision-makers will correctly and effectively employ the richer ensemble information. Although the present study is not about valuing ensemble forecasting, the insights into the way planners apply the uncertainty in environmental information may prove useful to Navy and Air Force weather organizations.

3.2 Experiment participants

Twenty-four students in the Graduate School of Operational and Information Sciences, at the Naval Postgraduate School (NPS), Monterey, CA, served as experimental participants during a ten-hour experiment conducted 28 February-11 March 2011. Their mean age was 33.4 years; services represented were Navy, 17, Marine Corps, 6, and Army, 1. Participants' rank ranged from 0-2 to 0-6. Six teams of four players participated in the experiment, with each playing the role of a Future Operations (FOPS) planning team under one of three experimental conditions.

3.3 Experiment overview

Similar to previous A2C2 experiments (Hutchins et al., 2009; 2010), teams were comprised of four players operating with a shared goal while working together over four two-hour sessions. Within the team, each player was assigned responsibility for a region (Area A or Area B) and planning period. In the four-day scenario, the current day was always designated as T; the plan for the next day was designated T+1; and, the plan 48 hours out was the T+2 plan. Players worked on plans in a rolling horizon, so that the current T+2 plan would be the starting point for T+1 plan developed tomorrow, ideally reflected in within-team coordination between the T+1 and T+2 planners for Areas A and B. Each player was given a weather forecast consistent with their area of responsibility within the theater, and with their planning period. The richness of this forecast information was varied by experimental condition.

Each player was also presented with a set of tasks to accomplish in their respective areas of responsibility and within the four play sessions. These tasks were oriented to the geography of the theater (Areas A and B), so that a team-within-team structure was enforced, with two players working on Area A tasks for T+1 and T+2, and similarly two players working on Area B tasks. All tasks were accomplished with shared (and limited) resources among the team members, which drove collaboration across geographic areas between A and B teams. Inside the team-within-team structure, the shared tasks for an area required coordination between T+1 and T+2 players. For example, the Area B team was given a set of tasks to accomplish within the four sessions, though in each session the T+2 player needed to know how much the T+1 player could accomplish in that planning period prior to building the T+2 plan.

Each player was given a weather forecast as a starting point for his planning work during the session, consistent with his area of responsibility (A or B), planning period (T+1 or T+2), and the team experimental condition in information richness. The experimental scenario was constructed in such a way as to create significant weather impacts and changes over the four sessions (Figure 1), though these events were phased so that in general if weather were good in one area it would be poor in the other area. While this weather phasing may seem contrived for the purpose of the experiment, it is worth noting that in large areas of responsibility (for example, in US Central Command and Southwest Asia) this is consistent with the observed weather pattern over a two to five-day period.

We hypothesize that planning teams with richer information will plan more effectively. In response to forecast weather conditions, we expect teams to (1) defer tasks until later in the period, when the weather is presently bad but forecast to improve; (2) accelerate tasks beyond normal workload, when the weather is presently good but forecast to deteriorate; and, (3) coordinate labor among areas of responsibilities so that tasks in fair weather regions can receive more attention to more effectively use resources within a 24-hour period in the experiment scenario. The experimental scenario was designed to clearly give players this trade space, while the dependent measures were intended to evaluate how well players operated in this trade space given different levels of information richness.

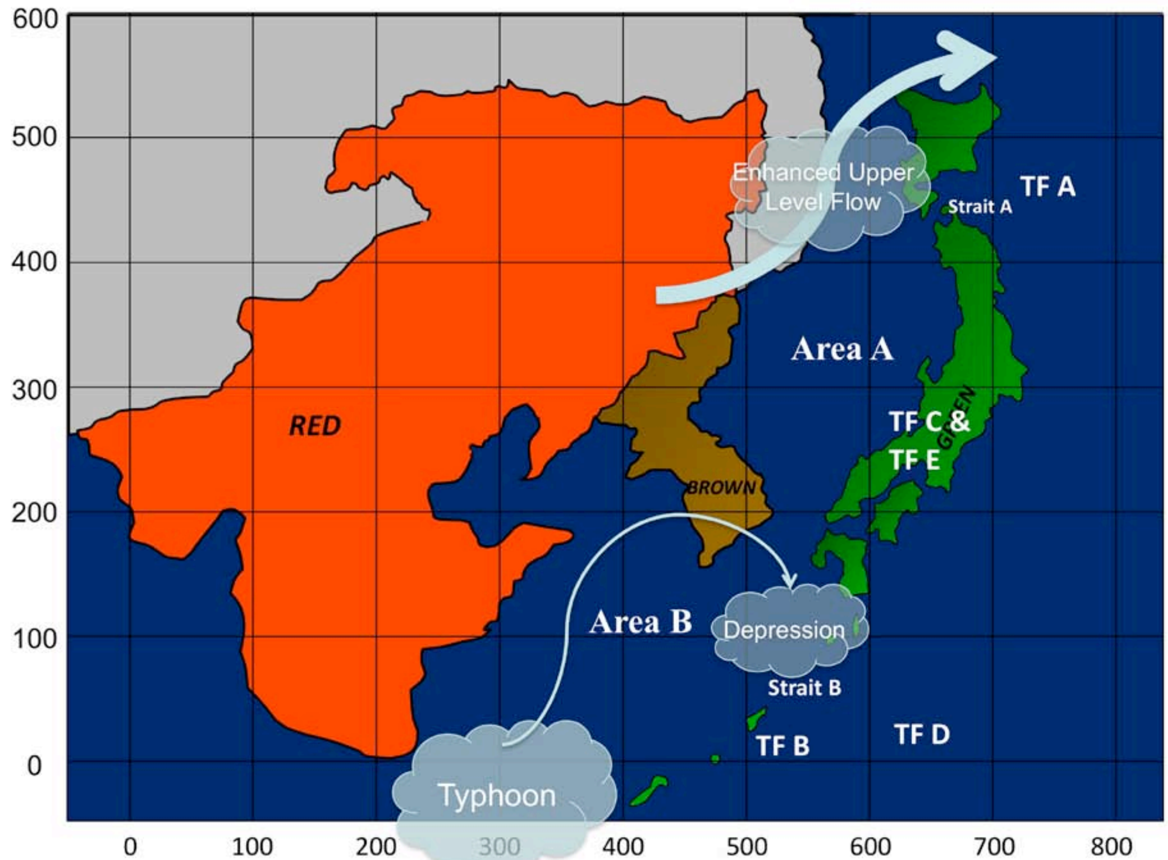


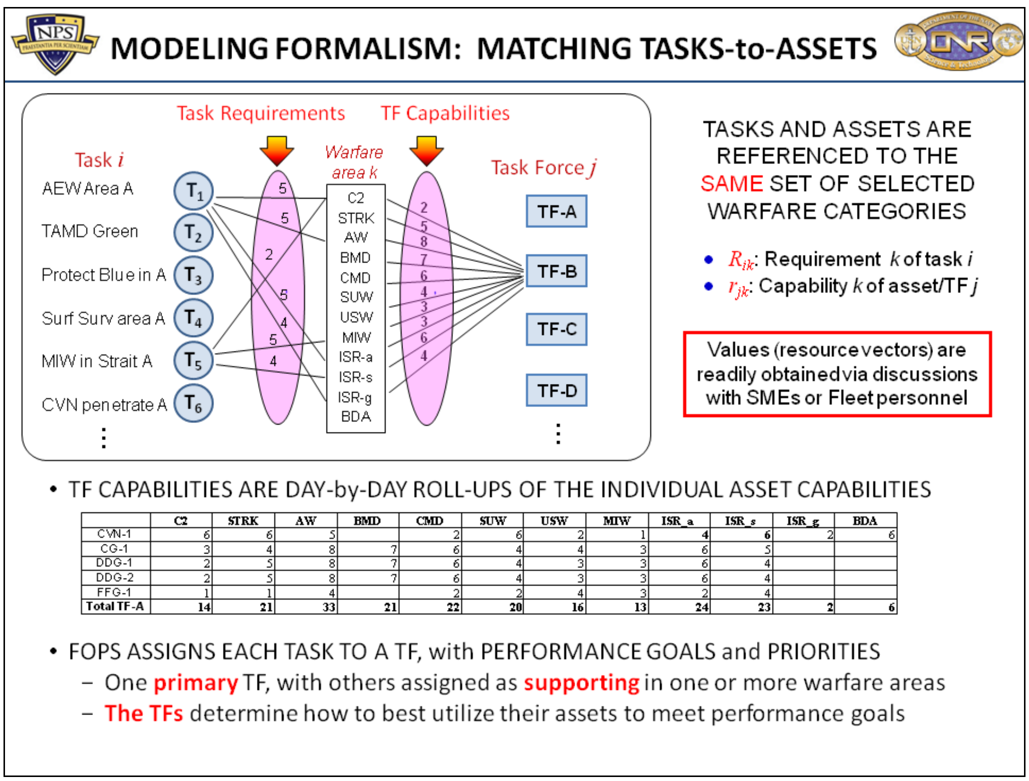
Figure 1. The experiment scenario was crafted over imaginary geography, with a challenging weather scenario affecting the tasks in Areas A and B. Task forces (TF A through E) are indicated as well.

3.4 Experiment design: Player planning tasks

All players completed their experimental work using a computer-mediated future operations (FOPS) planning tool developed at the UCONN (Mandal et al., 2010; Han et al., 2010). This tool served as the primary means of data collection during the experiment; all of the players' deliberations and final decisions were recorded for post-experiment analysis. Equally important, this tool served as a fundamental means to operationalize the FOPS planning process. In modeling the work of planners at the operational level, players made assignments at the task force level, while considering the tactical requirements for each assigned task and the tactical capabilities of each task force. The modeling formalism used to match assets to tasks is depicted in Figure 2.

Within the FOPS tool, players used the Assignment page (Figure 3) to reflect their chosen assignments of task forces to tasks. In game play, assigning a task involved: (a) selecting a primary task force, and (b) selecting up to two supporting task forces if needed to meet the warfare requirements of a task (Figure 3). For example, in Figure 3, the player is working on task TA02 (conduct intelligence, surveillance and reconnaissance on ground targets in Area A) that requires 32 units of the capability ISR_g, 16 units of capability C2 and 16 units of capability ISR_s. The player has selected Task Force A (TF-A) as the primary task force for this task.

In addition to requesting a task force (or task forces) for the task at hand, the planner may adjust the desired percent complete (Figure 3, dashed oval at top center) to better fit the available resources. The intent was for players to meet their desired percent complete on a task within the 24-hour (T+1) planning cycle without requesting more effort than available (because of weather or task force limitations) and without underestimating the possible effort achievable with the forecast weather conditions. Dependent measures in the experiment captured this asset efficiency as one measure of effective planning.



	C2	STRK	AW	BMD	CMD	SUW	USW	MIW	ISR_a	ISR_s	ISR_g	BDA
CVN-1	6	6	5		2	6	2	1	4	6	2	6
CG-1	3	4	8	7	6	4	4	3	6	5		
DDG-1	2	5	8	7	6	4	3	3	6	4		
DDG-2	2	5	8	7	6	4	3	3	6	4		
FFG-1	1	1	4		2	2	4	3	2	4		
Total TF-A	**14**	**21**	**33**	**21**	**22**	**20**	**16**	**13**	**24**	**23**	**2**	**6**

Figure 2. In the modeling formalism within this experiment, players assign Task Forces (comprised of multiple warfare areas or assets) to tasks. A limit on the number of tasks to which a Task Force may be assigned helps to create tension in the players' trade space.

The task force composition and warfare capabilities were developed as part of the experiment scenario. Relative numbers (e.g. 50 units of C2) were developed from discussions with subject matter experts. Clearly, resource scarcity could be engineered from this perspective, giving players only a few options to arrive at a "correct" solution; this approach was used in previous experiments (Hutchins et al., 2010). In order to drive planning teams to employ environmental information in their deliberations, however, the resources as assigned initially were chosen to make the four-session scenario relatively easy to complete under perfect weather. Resource scarcity, then, was almost solely a function of weather impacts in Areas A and B (Figure 1).

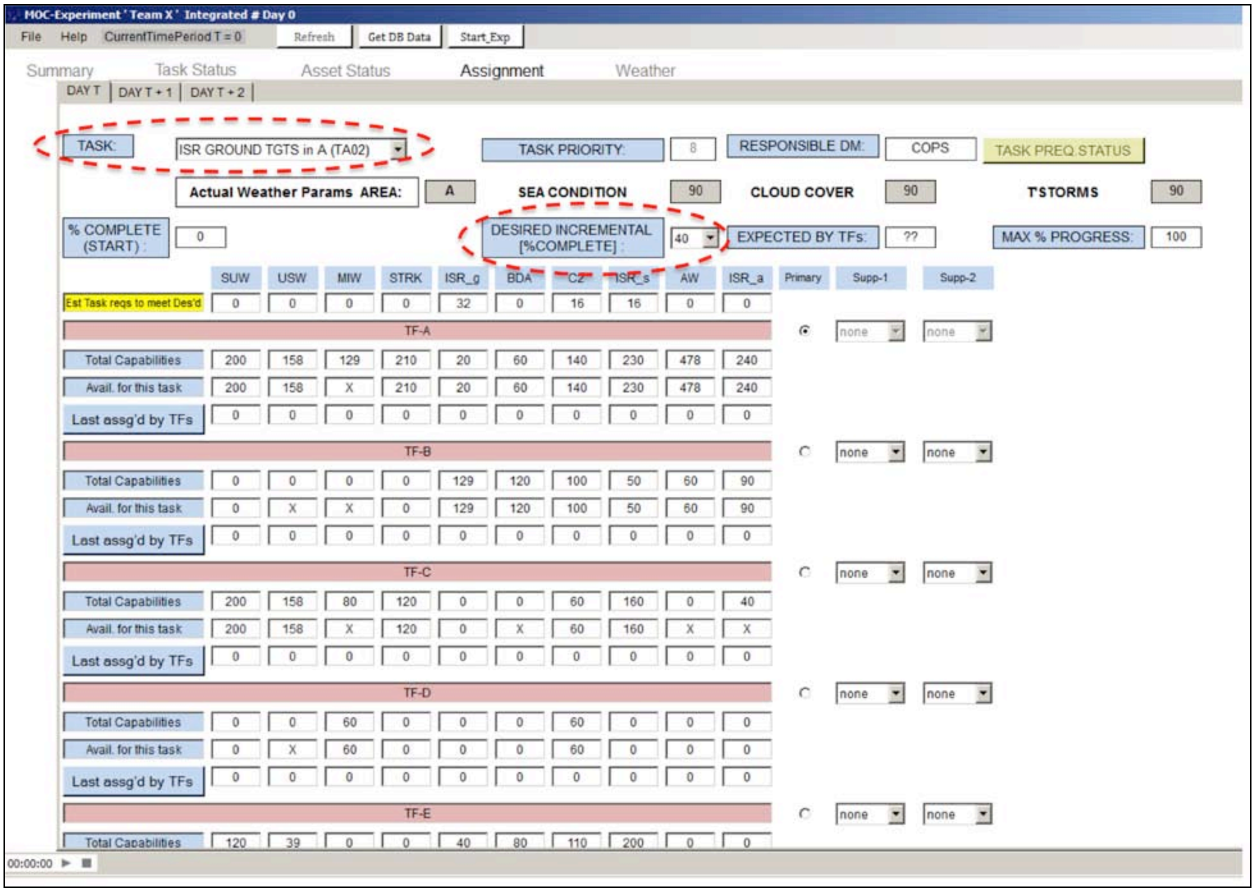


Figure 3. The Assignment page from the FOPS planning tool shows the task being worked (dashed oval in the upper left of the figure) and the task force assignments. The player's selected desired percent complete is indicated by the dashed oval at top center.

3.5 Experiment design: Operationalizing weather impacts

The Weather page (Figure 4) and Assignment page (Figure 3) were used in concert by planning teams to coordinate on tasks and resources within a planning session. The Weather page presented both forecast weather (controlled by the scenario) and a what-if assumed weather condition (Figure 4). To simplify the environmental information within the experiment, the players were presented with only three weather parameters: sea condition, cloud cover, and thunderstorms. These weather parameters were each expressed on a scale from 0 to 100, where 100 indicates perfect weather with no impact to warfare areas (Figure 5). Although players were not given the weather curves for the scenario warfare areas (Figure 5), the range of expected forecast impacts were indicated with a simple Assumed +5/-5 evaluation (Figure 4, center). This range is based on the players' choice of assumed weather and shifted as players engaged in what-if scenarios around the forecast weather conditions.

Given a task in a bad-weather region, players could opt to defer the task to later in the four sessions, or attempt task completion by applying additional assets in warfare areas affected by the weather conditions. In the limited assets available to the team, however, this would require coordination among the Area A and B teams-within-team.

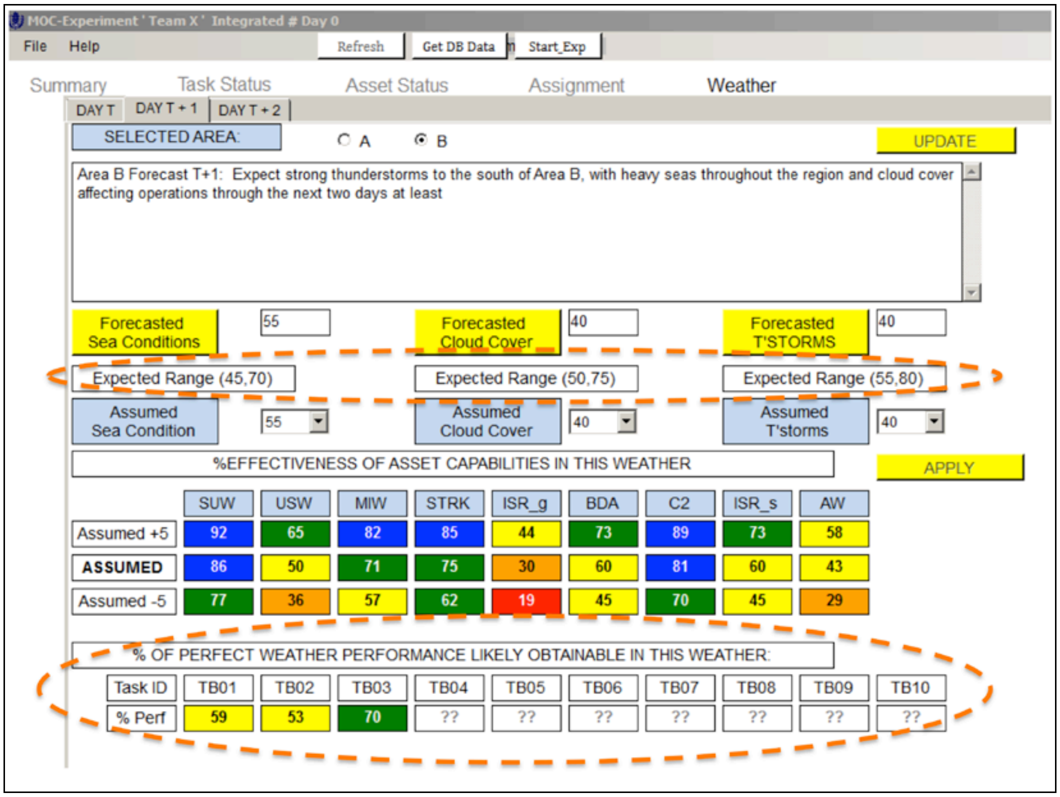


Figure 4. The Weather page provides planners with forecast weather conditions, along with additional uncertainty and integrative information per the experimental condition. This example is for Level III, with explicit uncertainty bounds (center dashed oval) and integrative information regarding the tasks (bottom dashed oval).

The independent variables were operationalized on the Weather page, with low-richness players operating with no forecast bounds or integrative mission information (indicated in Figure 4). With moderate information richness, players were given the explicit uncertainty bounds (dashed oval, center of Figure 4), though planning teams still needed to use the what-if tool to examine the possible weather impacts. In the high-richness condition, planning teams were given both uncertainty bounds and integrative information (Figure 4). Because the what-if tool was available in all experimental conditions, the number of times players employed this tool represented another useful dependent measure to assess players' awareness and application of information about the natural environment.

3.6 Experiment execution

The experiment was conducted in four two-hour time blocks, spread over a two-week period. The MOC Director (played by a confederate) presented the Commander's Update Briefing at the beginning of each experimental session. This briefing reviewed the scenario with players and updated teams on any changes in time-phased force deployment data (TPFDD) and resources available to the players. The MOC

Director also reviewed the task graph with the players to give an informal assessment of team progress on assigned tasks. A sample task graph is depicted in Figure 6.

Block 0 consisted of an introduction to the experiment, including a brief on the mission, and initial training in their team roles. Players also practiced with the computer-mediated FOPS tool.

Block 1 was the first full session where players were given the initial state of the scenario and tasked with building plans for Block 2 (T+1) and Block 3 (T+2). FOPS obtained updated information throughout the session to include updates to the weather forecasts for T+1 and T+2. The FOPS planning update for T+1 and T+2 (Blocks 2 and 3) was briefed to the MOC Director and submitted as a new plan at the end of this block.

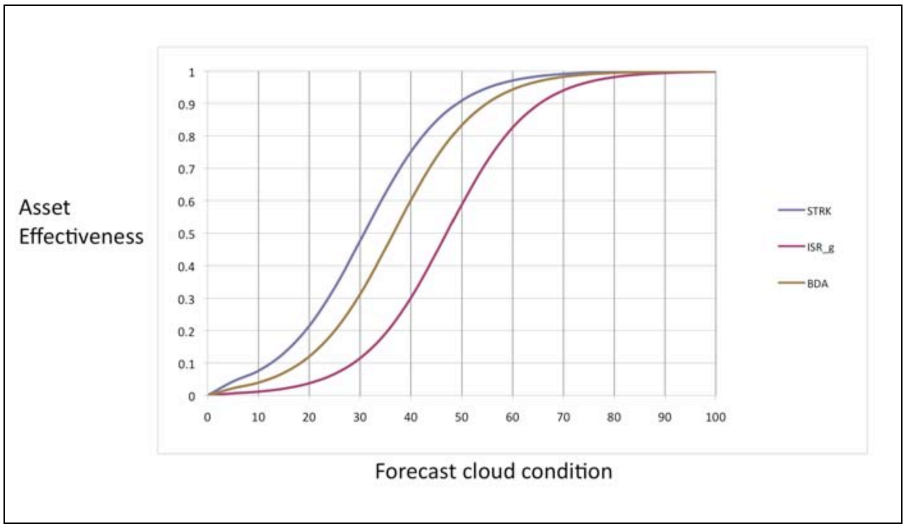


Figure 5. Response curves relating forecast cloud condition (on a scale of 0 to 100) to expected asset effectiveness (on a scale of 0 to 1). For the purpose of the experiment, the forecast condition closer to 100 means better weather and less impact to asset effectiveness. In this case, cloud cover effects strike (STRK), ground reconnaissance (ISR_g), and battle damage assessment (BDA).

Block 2 was the second session, beginning with the implementation of the FOPS team plan produced in Block 1. The FOPS team created a plan for Block 3 using their Block 1 plan for T+2 as a first guess. This team also created the Block 4 (T+2) plan based on their updated Block 3 (T+1) plan and previous progress.

Block 3 was the third session, and FOPS teams implemented the Block 2 plan for T+1, and used the Block 2 T+2 plan as the starting point for the T+1 (Block 4) plan. In this session the FOPS team also created a T+2 (Block 5) plan, although no such block was executed in this experiment. A briefing with the MOC Director was held at the end of Block 3.

Block 4 was the final session of the experiment. This session implemented the Block 3 plan for T+1, and player teams used the Block 3 T+2 plan as the starting point for the T+1 (Block 5) plan. A T+2 (Block 6) plan was also created based on the actual progress to date, and the expected progress of the T+1 (Block 5) plan. Neither of these plans for Block 5 or 6 (T+1 or T+2) is executed in the experiment scenario.

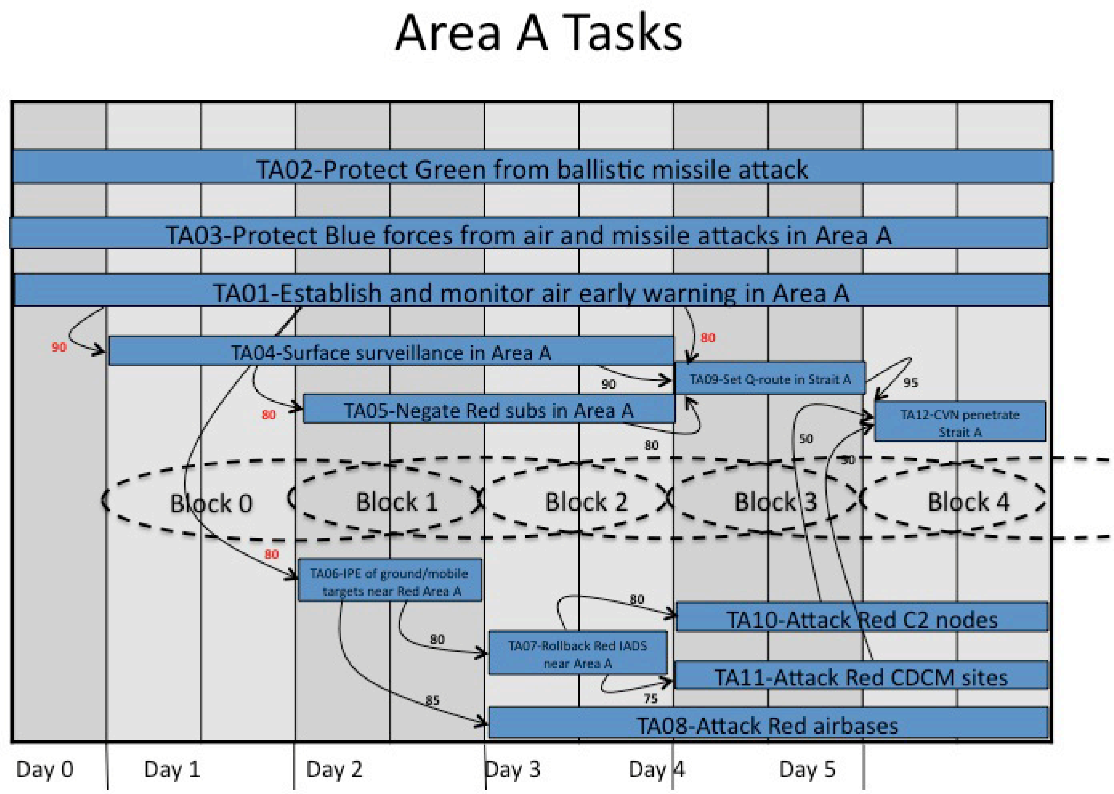


Figure 6. Sample task graph for Area A players. Dashed ellipses indicate the planning window for block and day, where Day 0 plans Block 0. Note that Block 0 was a training session and all players were given a notional plan at the end of the session so that all teams started Block 1 with the same initial conditions.

Within the four sessions, there are three opportunities to examine overlapping plans as a dependent measure. For example, the T+2 plan for Block 2 becomes the initial guess for the Block 3 T+1 plan. This plan volatility provides a useful insight into the effectiveness of the planning process for the team.

In addition to dependent measures collected within the FOPS planning tool, players were also asked to brief their plans, to include their rationale for task assignments. This permitted subjective data collection with additional insight into how the players across experimental conditions chose to integrate weather into their planning processes.

3.7 Independent variables

Viewed as an information processing entity, the organization seeks information to reduce uncertainty, with adjustments in collaborative and coordinating structures tuned to better accomplish this goal (Galbraith, 1974). In the present work, we do not modify the collaborative structure but instead examine the changes in media richness (Daft & Lengel, 1986) for effect on the information processing of planning teams within a modeled Maritime Operations Center.

The levels of the independent variable as originally conceived separated content and structure of the weather forecast products delivered into a classic 2X2 design (Table 1).

		Information Content (a, b)	
		Forecast Only	*Forecast with Explicit Uncertainty*
Information Structure (I, II)	*Automated Products*	**Level Ia**	**Level Ib**
	Automated Products with Human Expertise	**Level IIa**	**Level IIb**

Table 1. Original experimental concept for levels of the independent variables.

In this approach, Levels Ia and IIa represent much of the current practice in the Navy, Air Force and National Weather Service (Morss, 2010). Operationally, most atmosphere and ocean products are presented as deterministic forecasts with implicit uncertainty (Ia). This uncertainty is often clarified by additional information from experienced human forecasters (IIa). Analysis of these levels would yield some insight into the substantive contribution by human forecasters to the planning process.

Manipulation of information content would make explicit the uncertainty attached to the weather forecast (Levels Ib and IIb). This could be examined with (IIb) and without (Ib) human clarification, though the current literature suggests that correct interpretation of explicit uncertainty in environmental products would require some further expertise by weather specialists (Doswell, 2004). While this independent variable may still be interesting, it is possible that Level IIb could muddle the notion of the weather specialist bringing to bear meteorological expertise *and* mission-focused expertise. Consider, for example, a case where we know that a particular aviation platform requires surface winds below 10 nautical miles per hour (knots, or kts) to take off, and the forecast for tomorrow is 8 kts with a standard deviation of 3 kts. Our meteorological expert could offer that (1) the forecast cross winds will be 8 kts with a good chance of gusts to 11 kts; or, (2) forecast cross winds tomorrow have a good chance of violating the take-off threshold for our aviation platform, with a mean of 8 kts but a strong possibility of 11 kts in the period. Cross winds, here, refers to the wind component oriented 90 degrees across the runway. We assert that this additional, integrative information relating weather to mission impact is itself an interesting independent variable, though a 2X3 experimental design becomes difficult to manage.

To better meet our experimental objectives we chose to examine both aspects of the weather specialists' contribution to information processing using a compromise design. We transformed the 2X2 structure into a 1X3 approach, collapsing information structure and content into a single independent variable, information richness, as depicted in Table 2.

	Level I	**Level II**	**Level III**
Information Richness	*Forecast only, no additional weather or mission information*	*Forecast with explicit uncertainty information*	*Forecast with explicit uncertainty and integrating mission information*

Table 2. Adjusted design for the 2011 experiment.

In this approach, teams in Level I were presented with a simple weather forecast relevant to their planning activities but without explicit uncertainty nor with integrating mission-impact information (human expertise); this condition represents the control for Levels II and III. In Level II, the weather forecasts were provided to teams with explicit uncertainty information (e.g., winds tomorrow will be 8 kts plus or minus 3 kts), with additional expertise available *only* with respect to the forecast. In Level III, weather forecasts with uncertainty and integrating, mission-specific impact assessments were provided to teams for their planning activities.

In terms of information richness, we characterize Level I as low and Level III as high. While we expect team performance to improve from Level I to Level III, the magnitude of these differences, and the distinctions between Levels II and III, are the focus of our experiment and analysis. We next discuss the experimental measures used to examine these differences.

3.8 Dependent measures

As information richness increases from low to high, we hypothesize that planning teams given richer environmental information will perform better. The comprehensive data collection within the FOPS planning tool permitted examination of several dependent variables to quantify performance. Broadly, these can be thought of as: completion measures; efficiency measures; and, counting measures.

Completion measures were used to evaluate team performance in completing the assigned tasks within their task graph (see, for example, Figure 6). The overall task completion at the end of the four sessions offered an absolute completion measure and an overall means of comparing teams across the levels of the independent variable. We hypothesize that teams in Level III (integrative weather and mission impact information) will outperform teams in Levels II and I.

The task progress per session represented an excellent mechanism for immediate team feedback. That is, at the beginning of the session, players were briefed on their relative progress in the past session based on observed weather conditions and task force assignments. Because the players' trade space spans several sessions, this per-session measure is likely not as useful for evaluating team performance.

Efficiency measures provided insight into how well a team applied assets to tasks. Asset efficiency is maximized when assets are applied to tasks in relatively good weather; asset efficiency is minimized when players attempted to accomplish tasks in bad weather areas. This measure indicates how well a team adjusted to bad weather by shifting work from one area to the other. We hypothesize that teams in Level II and III will make more efficient use of assets than teams in Level I. We expect, too, that Level III will be more efficient than Level II.

Counting measures reflect the number of times something happened in game play. In particular, we assessed the players' belief in the weather forecast by examining the number of times the player used the what-if tool to change weather parameters. Each of the four members of a planning team had access to the weather what-if mechanism for their planning window (T+1 or T+2) and area (Area A or B), so this measure could be evaluated at both the level of the player and the team. At the team unit of analysis, we hypothesize that players with richer information will spend more time exploring weather and mission impacts, while players with less rich information will be more likely to simply accept the forecast weather conditions as given.

Another useful counting measure is the change of primary task force between overlapping plans. Blocks 2, 3, and 4 offered overlapping plans for the T+1 player, who started with the T+2 plan from the previous day as an initial plan. The changes between the final T+1 plan and the initial (previous day's T+2) plan represent a measure of volatility and give some insight into the robustness of the team planning process. Less volatility represents a more robust plan, and we hypothesize that teams with richer information will construct less volatile (more robust) plans.

3.9 Summary

Does providing more information to planners improve the planning process? Using a scenario driven by environmental impacts and modeled on MOC planning tasks, we investigated this question with a computer-mediated team-in-the-loop experiment. Information richness was manipulated as the independent variable across three levels ranging from low to high. Dependent measures of the planning process include assessments of asset efficiency; absolute and complete progress in assigned tasks; plan volatility between overlapping T+1 and T+2 plans; and subjective measures from post-session player briefings and surveys. We next present an analysis of the experimental outcomes.

4. Results

The experiment was conducted in late February and early March 2011 at the Naval Postgraduate School (NPS), Monterey, California. In this section we summarize the experiment execution and highlight the most significant results.

4.1 Participant demographics and team composition

Participants were recruited from the graduate student population at the NPS. These 24 subjects reflected a mean age of 32 (standard deviation ± 2.7 years), and had, on average, 11 years of military service. About half the participants had participated in previous A2C2 experiments. The subjects were grouped in to teams of four, with six teams total, or two per each level of the independent variable (Table 3).

	Level I	**Level II**	**Level III**
Experimental Groups	*Groups A and D*	*Groups B and E*	*Groups C and F*

Table 3. Experimental groups and team assignment.

We sought to maximize random assignment among the teams to balance for age, gender, military experience and other exogenous variables collected in the initial demographics. The strict requirement to meet student scheduling, though, meant that Groups C and F were somewhat younger (team average of 29 years) and consequently less experienced as military members (team average of 8 years).

4.2 Critical task completion measures

Each four-person experimental group was further sub-divided in to two two-person geographic teams, each with responsibility for about half of the area of responsibility (Figure 1). Each geographic team was given 10 tasks to complete within the four sessions of game play, with some tasks requiring careful coordination between the teams. An example task graph is given in Figure 6.

Of these 10 tasks, several were designed as "critical" tasks with significant prerequisites. These critical tasks were designed to force the experimental group to balance mission accomplishment with asset availability under expected weather conditions. The weather scenario was crafted intentionally to increase the player workload and require more intense deliberation in the completion of these tasks.

An overall completion score for all tasks was computed for each group, with scores ranging from 88% to 95%. This overall measure represents the mean completion rate for all tasks assigned (Table 4). To examine the performance on critical tasks, a deviation score was computed based on the incompleteness of the six critical tasks, reflected as a deviation from perfect (Table 4). This deviation, expressed as a negative number in Table 4, represents a direct measure of the group performance in critical, highly weather-dependent tasks, in that a smaller (less negative) deviation represents better performance.

	Groups	**Overall Completion Score**	**Composite Overall Completion**	**Critical Task Completion Deviation**	**Composite Critical Task Deviation**
Level I	A	93	91.5	-23	-30.5
	D	91		-38	
Level II	B	88	88.5	-58	-48.5
	E	89		-39	
Level III	C	95	94.5	-11	-15.5
	F	94		-20	

Table 4. Overall and critical task completion scores by group and level.

A one-way analysis of variance on the overall scores in Table 4 showed a significant difference among levels at $p < 0.02$, suggesting that the manipulation of information richness did indeed have an impact on team performance. In examining the critical task deviations, though, the differences among all levels of the independent variable drop to a critical level of $p < 0.11$. While still significant, this suggests that perhaps some groups, regardless of level, may have had some degree of difficulty in understanding and integrating critical weather information.

One oddity in these data is the relatively poor performance of the Level II groups on critical and overall task completion. These groups were presented with a weather forecast with a fixed confidence interval but no further integrating information. No demographic differences were apparent in these groups (B and E), and the consistency in poor performance of both groups in Level II suggests some inconsistency in this particular instantiation of the independent variable.

4.3 Efficiency measures

While the group measures of task completion give broad insight into team performance, an examination of plan efficiency among groups gives more subtle and a more useful perspective on the players' application and integration of environmental information. Plan efficiency was measured in post-experiment analysis by comparing the daily plans produced within groups against an ideal plan based on perfect knowledge of the weather. Monte Carlo simulations were used to identify the most asset-efficient task force assignment to tasks given the current disposition of forces and the current (true) weather conditions. Comparisons of player plans to these ideal plans showed scores from 92 to 100% (Table 5).

A one-way analysis of variance on the data in Table 5 shows significance at $p < 0.09$, and the direction of the relation is consistent with our experimental hypotheses. That is, groups with richer information appear to make more efficient use of resources in constructing their plans. One reasonable inference from these data is that players with richer information make better global assignments of assets to mission requirements, particularly when this information included integrated mission impact data connected to forecast weather and ocean conditions. Although the p-value around 0.09 suggests a less powerful result, we should keep in mind that the teams in Level III were both younger and less experienced than their counterparts in Levels II and I. Had we been able to more adequately randomize the groups we believe these efficiency measures would have shown an even more dramatic difference among levels.

	Groups	Overall Completion Score	Composite Overall Completion
Level I	A	92	93
	D	94	
Level II	B	99	98
	E	97	
Level III	C	100	99
	F	98	

Table 5. Efficiency measures of group performance.

4.4 Subjective measures

In addition to measures of task completion and efficiency, subjective measures were collected from the players at the end of each game session. These surveys were used to assess cognitive workload and self-assessment of performance by each of the groups.

Participants in the Level I groups were split roughly 50:50 when asked if they made changes from the forecast regarding the weather. About half stated they thought the weather would be better and they tended to be optimistic about the weather, while half indicated they tended to be pessimistic. Many indicated that there were not a lot of data points regarding the weather, consistent with this level of the independent variable.

Participants in the Level II groups demonstrated a consistent pattern of making changes to the assumed weather from the forecast. They indicated they made changes 91 percent of the time and did so for the following reasons: because the weather update indicated more favorable conditions; because the weather forecast seemed overly conservative when compared to the available range; and because to increase performance on task prerequisites that needed to be accomplished to be better positioned for follow-on planning for tasks. This suggests that given explicit uncertainty information, player planners did try to make good use of this information.

Participants in the Level III groups indicated they made changes from the forecast 79% of the time. Their reasons for making changes were based on updates that showed better or worse than expected weather; recommendations from the METOC officer (an automated confederate); distance between ranges and the pattern of the weather; and to achieve better performance in terms of time to complete tasks. This also suggests that given explicit uncertainty information as well as mission-impact information (e.g., thunderstorms at 48 hours may affect air warfare) player planners will include this information in their decision-making processes.

Players were asked to rate on a scale of 0 to 100 their perceived mental effort; overall effort; time pressure; frustration; and performance. The composite means for all players in a group appear in Table 6; players were surveyed a total of four times over the experiment.

Simple one-way analyses of variance on these factors highlight several interesting features of the experiment. Both perceived mental effort and perceived overall effort show little distinction among levels of the independent variable; a reasonable inference from this evidence is that the weather scenario as designed did not necessarily favor the Level II and III groups. The perceived time pressure, though, did show some significant difference among groups ($p < 0.13$) though the Level III groups appeared to feel the most pressure. We speculate that the relative youth and inexperience of this group may have contributed to a keener sense of the time pressure in the experiment. This may also have been simple systematic error in measurement, though, as the perceived levels of frustration ($p < 0.57$) and performance ($p < 0.73$) showed little distinction among the groups.

	Group	Mental Effort	Overall Effort	Time Pressure	Frustration	Performance
Level I	A	46.7	40.7	24.0	29.7	87.0
	D	39.5	34.0	25.0	31.5	43.0
Level II	B	40.7	38.3	15.0	33.3	45.0
	E	48.0	42.7	27.3	22.3	93.7
Level III	C	45.3	43.3	36.6	45.3	65.0
	F	44.0	39.3	34.4	28.6	87.0

Table 6. Self-reported measures from post-session surveys.

A multiple analysis of variance on these data (Table 6) shows several expected interactions among measures. In particular, perceived frustration and perceived performance were significantly negatively correlated ($r = -0.71$). Frustration, however, appeared to be relatively uncorrelated with perceived overall effort ($r = 0.03$). We speculate that frustration, as a measure, was more indicative of player comfort with the computer-mediated simulation rather than with information delivered under different experimental conditions.

4.5 Summary

Direct measures of group performance support the working hypothesis that teams given richer information produce better and more efficient plans. The subjective measures collected during the experiment suggest that the performance differences were attributable to real differences among the levels of the independent variable, and not simple differences among teams and team members. We next conclude our discussion and present several ideas for future research.

5. Discussion

Does providing more information to planners improve the planning process? In this study we have investigated planning under uncertainty, manipulating information richness as the independent variable. Human subjects, playing the role of Maritime Operations Center planners, participated in an eight-hour computer-mediated game in a scenario where successful mission accomplishment required players evaluate and integrate expected weather information. Experimental results suggest that groups operating with richer information (weather forecasts with explicit uncertainty and integrated with mission impacts) significantly outperform planning teams operating with less rich information (simple deterministic forecasts).

Recent trends in both the Navy and Air Force provide significant motivation for this study. The increasing environmental sensitivity of weapons systems, particularly autonomous systems, places a higher premium on weather and ocean information in defense operations. The state of the science in atmospheric and ocean prediction requires a blend of human and automated processing to create actionable intelligence from imperfect numerical forecasts (Doswell, 2004; Lorenz, 1963). Although the focus of this study was on planning under uncertainty, we can think of the levels of the independent variable (information richness) as representing different combinations of human and automated processing. Future investigations could refine and more deeply examine the optimal mix of human and automation within the meteorological and oceanographic (METOC) communities.

In this study, planning under uncertainty focused on METOC data as the source of uncertain information. A future investigation might consider the more general case of highly uncertain intelligence in constructing operational plans. Sources of uncertainty in these studies might range from uncertainty in modeling the physical atmosphere and oceans, to estimates of enemy capabilities and intentions. Such a study would better inform applied or operational research in human-to-machine interfaces in command and control systems.

We have established in this work that providing richer information (explicit uncertainty) to planners enables planners to deliver more effective and efficient plans. We sought to model as closely as possible the planning processes inside the Maritime Operations Center (MOC). Future investigations would refine this model of the MOC, and extend this work into the Air Operations Center (AOC) and Joint Operations Center (JOC). Among the operational forces today, the proliferation of information (command and control) systems appears to put a wealth of data in the hands of planners, though not all of these data may be immediately useful or relevant. In this work, we have sought to better understand how these data become information in the human process of command and control with the addition of explicit uncertainty bounds on these data. We expect these results will inform further studies to improve understanding of the best mix of human and automated processing that leads to better performing teams and, ultimately, more effective planning.

Acknowledgements

This research was accomplished with support from the Office of Naval Research and under the project Adaptability in Information Flow for C2. The authors gratefully acknowledge the many productive conversations with Dr. Eva Regnier (Naval Postgraduate School) and Dr. Jim Hansen (Naval Research Laboratory) regarding weather forecasting, risk and uncertainty.

References

Daft, R.L. and Lengel, R.H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science,* 32(5): 554-571.

Doswell, C.A. (2004). Weather Forecasting by Humans—Heuristics and Decision Making. *Weather and Forecasting*, 19: 1115-1126.

Galbraith, J.R. (1974). Organization Design—An Information Processing View. *Interfaces,* 4(3): 28-36.

Hutchins, S. G., Kemple, W. G., Kleinman, D. L., Miller, S., Pfeiffer, K., Weil, S., Horn, Z., Puglisi, M., and Entin, E. (2009). Maritime Headquarters with Maritime Operations Center: A Research Agenda for Experimentation. In *Proceedings of the 14th International Command and Control Research and Technology Symposium* (ICCRTS), June 2009. Washington, DC.

Hutchins, S. G., Kemple, W. G., Kleinman, D. L., Miller, S., Pfeiffer, K., Weil, S., Horn, Z., and Entin, E. (2010). Maritime Operations Center with Integrated and Isolated Planning Teams. In *Proceedings of the 15th International Command and Control Research & Technology Symposium* (ICCRTS), June 2010. Santa Monica, CA.

Lorenz, E.N. (1963) Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2):130-141.

Leutbecher, M. & Palmer, T.N. (2008). Ensemble Forecasting. *Journal of Computational Physics*, 227: 3515-3539.

Moore, R. A., Schermerhorn, J. H., Oonk, H. M., and Morrison, J. M. (2003). Understanding and Improving Knowledge Transactions in Command and Control. In *Proceedings of the 8th International Command and Control Research & Technology Symposium* (ICCRTS), June 2003, Washington, DC.

Morss, R.E., Lazo, J.K., and Demuth, J.L. (2010). Examining the use of weather forecasts in decision scenarios: results from a US survey with implications for uncertainty communication. *Meteorological Applications*, 17: 149-162.

Nobis, T.E., Kuchera, E.L., Rentschler, S.A., Rugg, S.A., Cunningham, J.G., Snyder, C. and Hacker, J.P. (2008). Towards the Development of an Operational Mesoscale Ensemble System for the DoD Using the WRF-ARW Model. In *Proceedings of the High Performance Computing Modernization Program (HPCMP) Users Group Conference*, July 2008, Seattle, WA.

Regnier, E.D. (2008a). Public Evacuation Decisions and Hurricane Track Uncertainty. *Management Science*, 54(1): 16-28.

Regnier, E.D. (2008b). Doing Something about the Weather. *Omega: The International Journal of Management Science*, 36(1): 22-32.

Richardson, L.F. (1922). *Weather Prediction by Numerical Process*. New York: Cambridge University Press.

Sestak, M., Bishop, C., Holt, T., Nachamkin, J., Chen, S. McLay, J., and Doyle, J. (2008). Dedicated High Performance Project Investment (DHPI) for the Fleet Numerical/Air Force Weather Agency – The Navy Side. In *Proceedings of the HPCMP Users Group Conference*, July 2008, Seattle, WA.

US Fleet Forces Command (2007). *USFF Maritime Headquarters with Maritime Operations Center Fleet Governing Board (MHQ with MOC FGB),* Norfolk, VA. 17 Aug 2007.

US Marine Corps (1996). Marine Corps Doctrinal Publication 6, *Command and Control*. Washington, DC: Government Printing Office.

NAVAL
POSTGRADUATE
SCHOOL

Planning with uncertain and ambiguous information: Command, control and the natural environment

Karl D. Pfeiffer, Susan G. Hutchins, David L. Kleinman,
Scot A. Miller, and William G. Kemple

June 2011

Background

A2C2 Research Program

- Adaptive Architectures for Command and Control (A2C2) has operated as a research program at Naval Postgraduate School for over 15 years
  - Integrates analytic modeling, human-in-the-loop experimentation and simulation in a research paradigm of model-test-model-experiment
  - Models and associated simulations define and guide the experiments, and results from experiments are then used to improve models.
- Over the past three years, A2C2 investigators have developed a multi-disciplinary research agenda to explore issues critical to the Maritime Operations Centers (MOC).

Background

Motivation

- Operational planners are often faced with constructing robust, effective plans using ambiguous information in a complex or evolving situation.
- Environmental information represents a particular challenge for the planner—weather and ocean forecasts carry significant uncertainty.
- Even with perfect knowledge, translating these conditions into mission impacts can be difficult--*delivering a more accurate weather forecast does not necessarily provide the planners with more useful information.*

"Listen, S-2," the colonel said, "I don't care about how many inches of rainfall to expect. I don't care about the percentage of lunar illumination. I don't want lots of facts and figures. Number one, I don't have time, and number two, they don't do me any good. What I need is to know what it all means."

—USMC Doctrinal Publication 6 *Command and Control* (1996)

Background

Weather and Uncertainty

- Weather prediction is appealing as a purely deterministic problem—but even the current state of the science shows limited skill beyond six days

- Every numerical weather prediction carries an inherent uncertainty.
  - Typically, the bounds of this uncertainty are unknown, though a skilled human weather forecaster can attach a reasonable estimate of the uncertainty to a forecast.
  - Although this uncertainty estimate may be qualitative the human forecaster nonetheless can make this uncertainty clearer to decision-makers.

- Ensemble forecasting is an explicit approach to evaluate model uncertainty, comparing several models over the same forecast period
  - The explicit means and variances derivable from ensemble output provide another means to make clear to decision-makers how much trust to place in a particular forecast
  - *Interpretation of these numbers, however, is not always clear to decision-makers*

Background

Weather and Decision-making

- No matter the skill with which we predict the natural environment, these predictions are of themselves little use to military planners.
- Decision-makers are largely concerned with when and to what degree their assets and capabilities will be affected by weather conditions
- For planners with a trade-space spanning days or weeks, decisions to proceed with, accelerate or delay operations are connected to the expected atmospheric and ocean conditions.
- Effective planning in this case is connected to *knowledge and exploitation* of the natural environment.

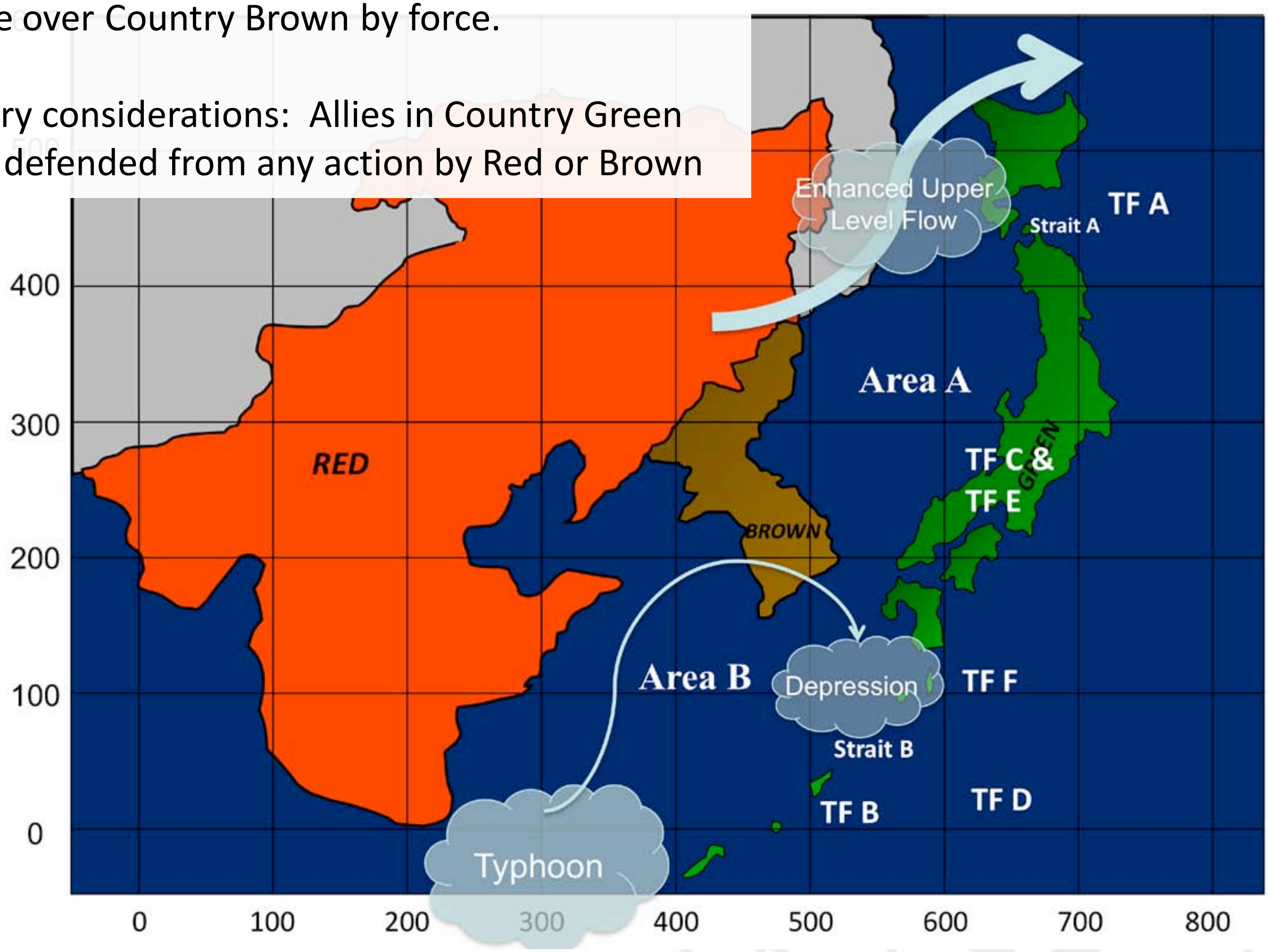
Research Questions

- We seek to better understand how organizations can employ perishable and uncertain information in the operational planning process.
  - In the face of inherently uncertain information, if we make this uncertainty more clear through human intercession, or explicit quantitative bounds, how will planners apply this information?
  - Given actionable mission impacts connected to this uncertain information, how will planners integrate and apply this information?
- In the context of the Maritime Operations Center, does providing richer information to planners lead to better planning?

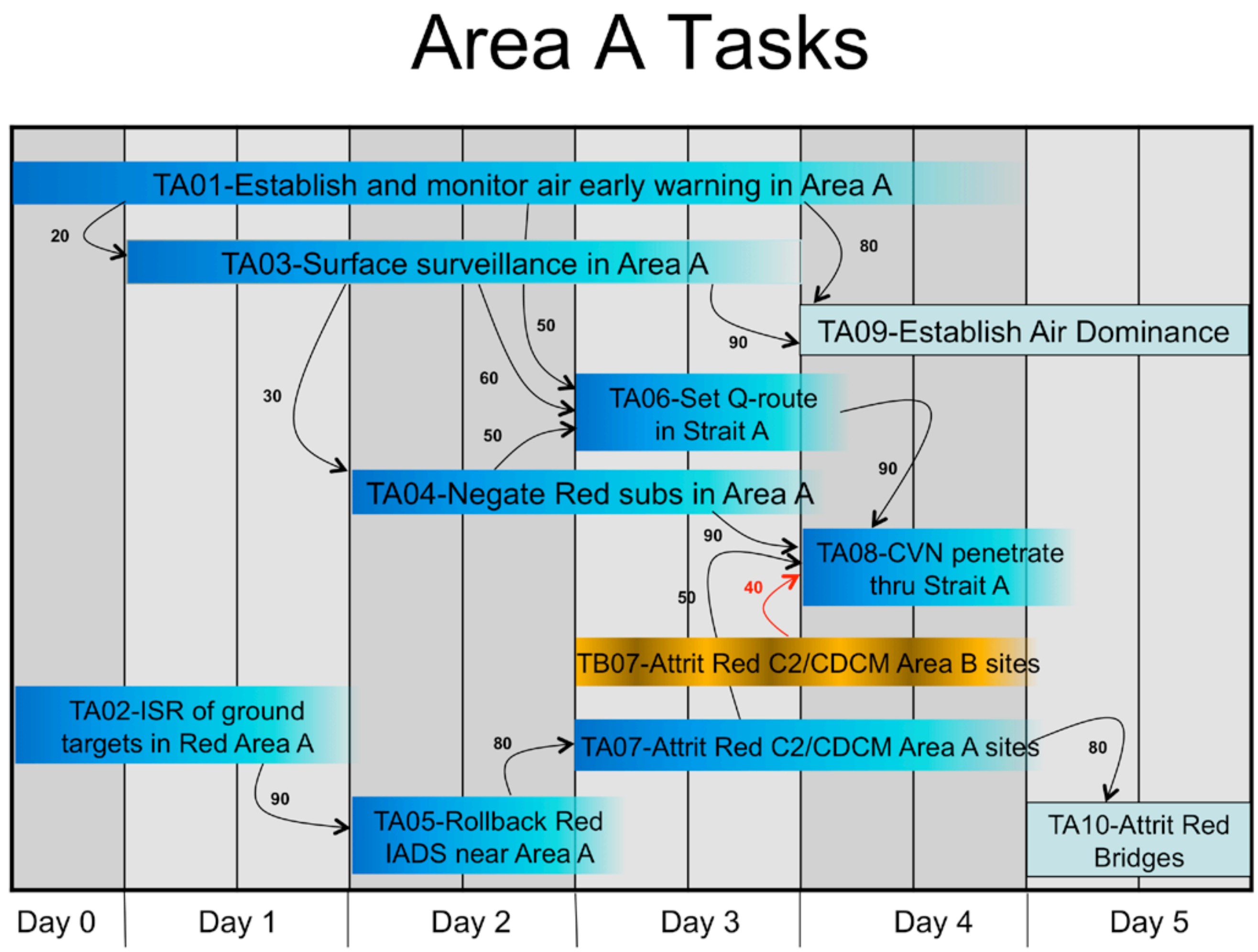
Experimental Design
Modeling Formalism: Matching Assets to Tasks

Objective: Break area denial that has been established by Country Red as it tries to extend its influence over Country Brown by force.

Secondary considerations: Allies in Country Green must be defended from any action by Red or Brown

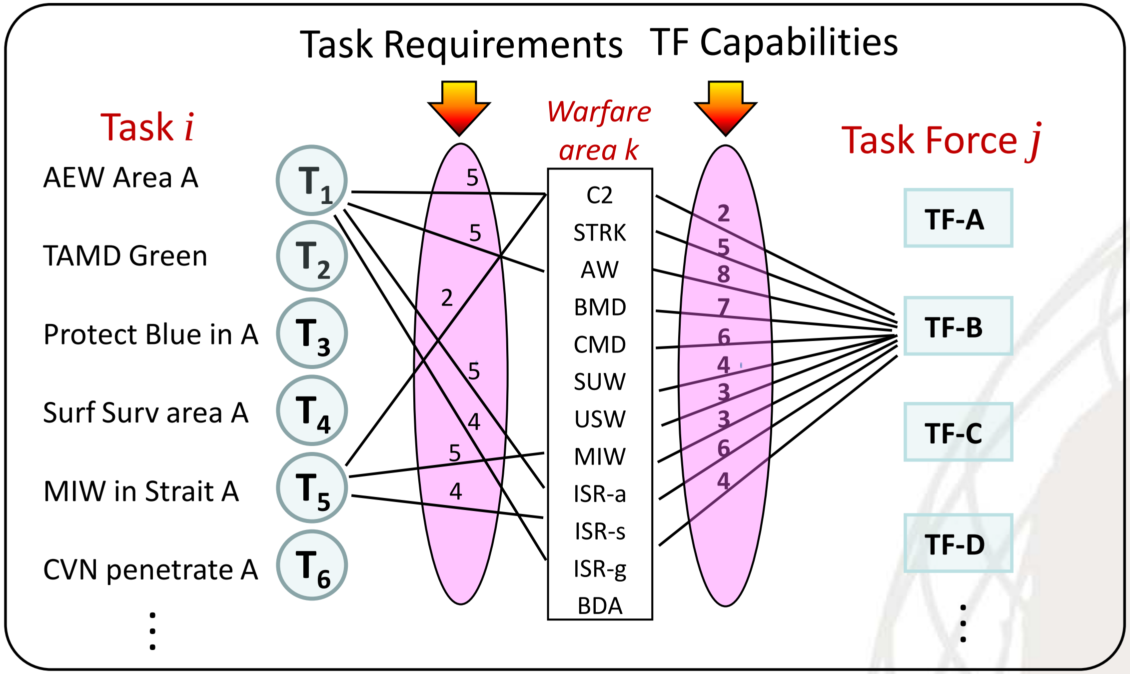


Experimental Method
Modeling Formalism: Matching Assets to Tasks



Experimental Method

Modeling Formalism: Matching Assets to Tasks



TASKS AND ASSETS ARE REFERENCED TO THE SAME SET OF SELECTED WARFARE CATEGORIES

- R_{ik}: Requirement *k* of task *i*
- r_{jk}: Capability *k* of asset/TF *j*

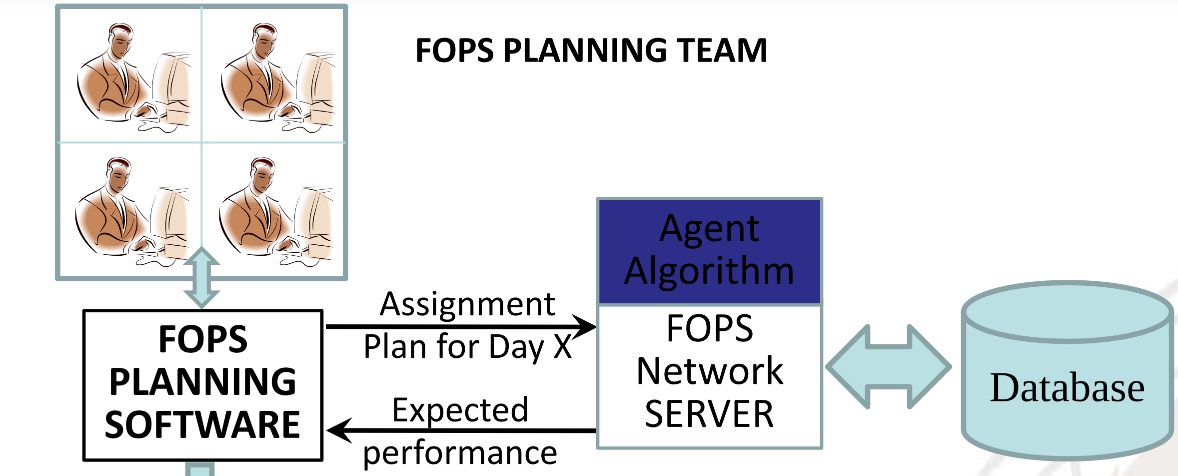
Values (resource vectors) are readily obtained via discussions with SMEs or Fleet personnel

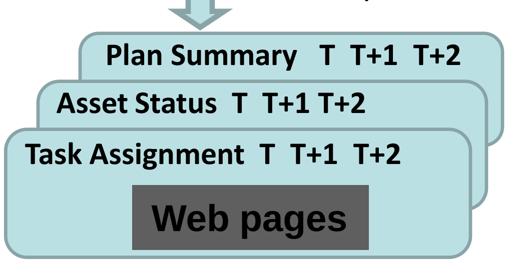
Future operations (FOPS) planners assign each task to a task force (TF)

- Planner assignments (requests) include *performance goals* and *priorities*
- One task force is designated as primary by the FOPS planner
- Planner may assign other TFs as supporting in one or more warfare area
- The task forces (computer agents) determine how to best use assets to meet goals

NAVAL POSTGRADUATE SCHOOL

Experimental Method
Laboratory Environment





- Static (task and asset) data
- Scenario information
- Dynamic information
- Automated data collection

Players produce a "plan" by assigning task forces to all active tasks on day T+1 and T+2, then UCONN agent algorithms allocate TF assets to best meet overall task performance goals

Each team member has a different planning responsibility:

	Area A	Area B
T+1	FOPS 1	FOPS 3
T+2	FOPS 2	FOPS 4

Experimental Design

Player Tasks

The team develops a plan by considering the critical task prerequisites; planned task start dates; and the weather forecasts for each area of responsibility.

Trial or working plans can be submitted to the task forces for review. This submission returns an expected performance for the given plan, based on assets available and the weather impacts to those assets. This evaluation is computed by the agent-based model.

The team then modifies assignments on those tasks not meeting desired criteria

When the team believes the plan is satisfactory, the plan is finalized so that:
T+1 plan => EXORD for tomorrow
T+2 plan => start for next T+1 plan

	Area A	**Area B**
T+1	FOPS 1	FOPS 3
T+2	FOPS 2	FOPS 4

Sequential Coordination

Parallel Coordination

Over the four 2-hour sessions, players *should* trade off assets from tasks in bad weather areas to tasks in good weather areas, and defer starting tasks from T+1 to T+2 (or longer, within commander’s guidance) if the weather is forecast to improve.

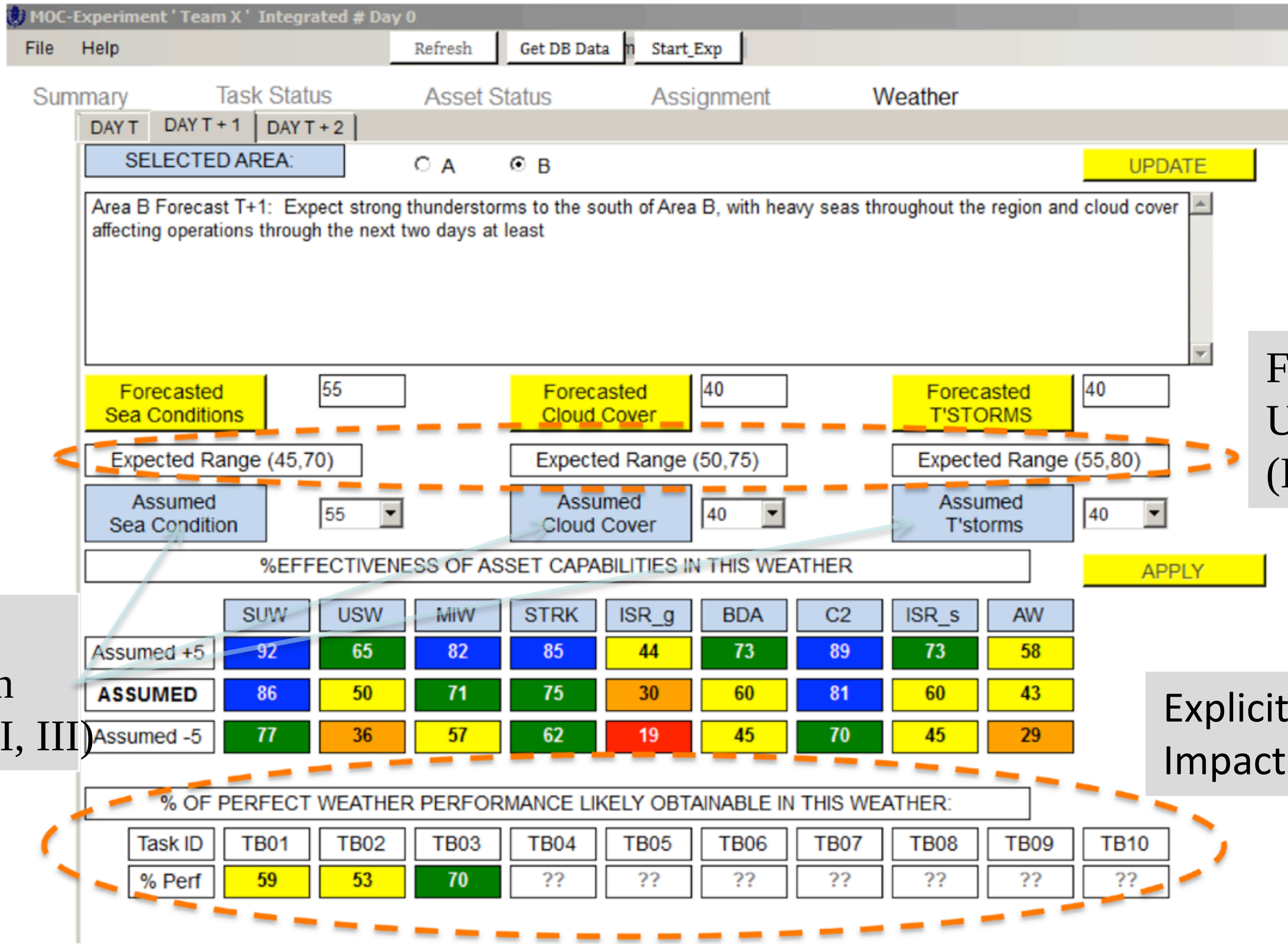
Experimental Design
Independent Variables

	Level I	Level II	Level III
Information Richness	*Forecast only, no additional weather or mission information*	*Forecast with explicit uncertainty information but no additional mission information*	*Forecast with explicit uncertainty and integrating mission information*

To better meet our experimental objectives we chose to examine both aspects of the weather specialists' contribution to information processing using a compromise design.

Experimental Design

Operationalizing the independent variable



Forecast Uncertainty (Level II, III)

What-if Evaluation (Level I, II, III)

Explicit Mission Impacts (Level III)

Experimental Design

Dependent Measures

- Broadly, we hypothesize that planning teams given richer environmental information will perform better. “Better” in this sense means that teams made better use of limited assets, and, ultimately, produced better plans

- Completion measures are used to evaluate team performance in completing the assigned tasks within their task graphs
  - We hypothesize that teams in Level III (integrative weather and mission impact information) will outperform teams in Levels I and II.

- Efficiency measures examine how well a team applied assets to tasks.
  - Indicates how well a team adjusted to bad weather by shifting work from between areas
  - We hypothesize that teams in Level II and III will make more efficient use of assets than teams in Level I. We expect, too, that Level III will be more efficient than Level II.

Experimental Results

Completion Measures

	Groups	Overall Completion Score	Composite Overall Completion	Critical Task Completion Deviation	Composite Critical Task Deviation
Level I	A	93	91.5	-23	-30.5
	D	91		-38	
Level II	B	88	88.5	-58	-48.5
	E	89		-39	
Level III	C	95	94.5	-11	-15.5
	F	94		-20	

- One-way ANOVA shows a significant difference at *p < 0.02*, suggesting that manipulation of information richness did impact team performance.
- For critical task deviations differences among all levels of drop to a critical level of *p < 0.11 … groups may have had trouble understanding the task graphs and the scenario critical path*

Experimental Results

Efficiency Measures

	Groups	Efficiency Score	Composite Efficiency Score
Level I	A	92	93
	D	94	
Level II	B	99	98
	E	97	
Level III	C	100	99
	F	98	

- One-way ANOVA shows significance at $p < 0.09$, and the direction of the relation is consistent with our experimental hypotheses.

- One inference from these data is that players with richer information make better global assignments of assets to mission requirements, particularly when this information included integrated mission impact data connected to forecast weather and ocean conditions.

Experimental Results

Survey Measures

	Group	Mental Effort	Overall Effort	Time Pressure	Frustration	Performance
Level I	A	46.7	40.7	24.0	29.7	87.0
	D	39.5	34.0	25.0	31.5	43.0
Level II	B	40.7	38.3	15.0	33.3	45.0
	E	48.0	42.7	27.3	22.3	93.7
Level III	C	45.3	43.3	36.6	45.3	65.0
	F	44.0	39.3	34.4	28.6	87.0

- MANOVA shows several interesting interactions among measures.
- Perceived frustration and perceived performance were significantly negatively correlated (r = -0.71)
- Frustration, however, appeared to be relatively uncorrelated with perceived overall effort (r = 0.03)
- We speculate that frustration, as a measure, was more indicative of player comfort with the computer-mediated simulation rather than with information delivered under different experimental conditions.

Conclusions

Operational relevance

- We expect that teams given richer information will engage in more effective planning and likely will produce a better plan … *so what?*

- Our design is intended to address the more useful question: *how much better do teams perform given richer uncertainty information?*

- This question is of significant operational relevance to both the Navy and Air Force, as there is a cost to keep humans deeply embedded in the forecast process, and a cost to produce explicit uncertainty bounds with numerical forecasts.

Conclusions

Operational relevance

- Within the DoD the current trend is to consolidate METOC personnel in centers located far from the forward edge of battle, and most often located in the CONUS. Support to deployed operations is then provided with online product delivery and reach back service to these centers.

- Both the Navy and the Air Force are examining the use of ensemble numerical weather prediction to improve operational forecasts and improve the explicit uncertainty information attached to these forecasts.

- For both services a lingering concern is whether decision-makers will correctly and effectively employ this richer information—insights from this study may prove useful to Navy and Air Force organizations shaping and re-shaping their decision-support and planning processes.

Conclusions

The way ahead

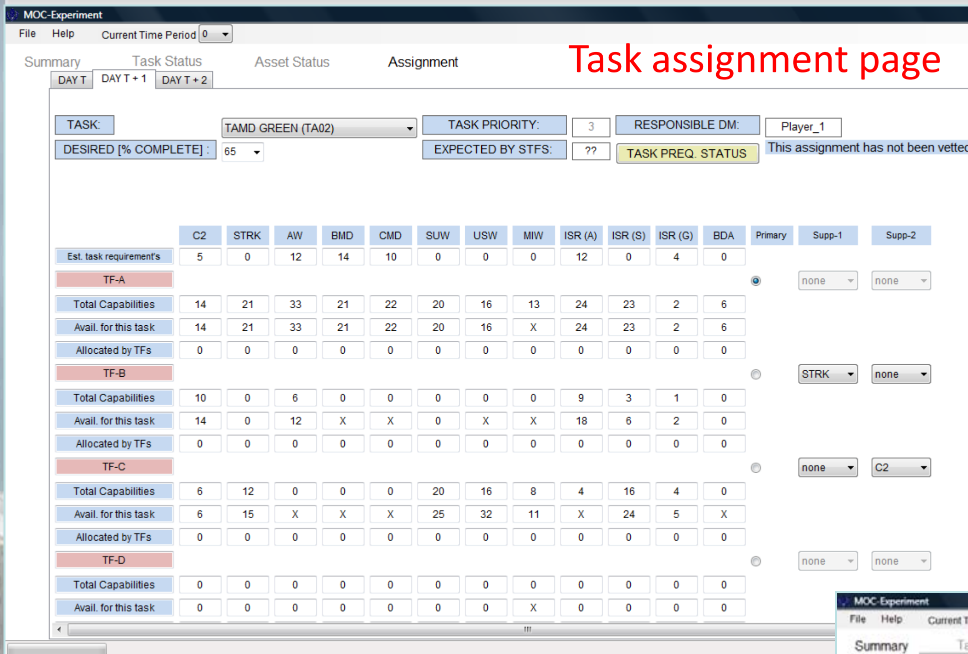
- This work motivates several avenues for future investigation, including:
  - Planning under high task uncertainty (e.g. reconnaissance, close air support, casualty evacuation) interacting with uncertainty in the natural environment
  - Measures of trust in weather forecasts and other intelligence products
  - Human factors in C2 systems: creating actionable intelligence from multiple sources
  - Active and passive deception detection in C2 planning systems

NAVAL
POSTGRADUATE
SCHOOL

Backup Slides

Experimental Design

Modeling Formalism: Matching Assets to Tasks



Plan is submitted to "TFs" for review

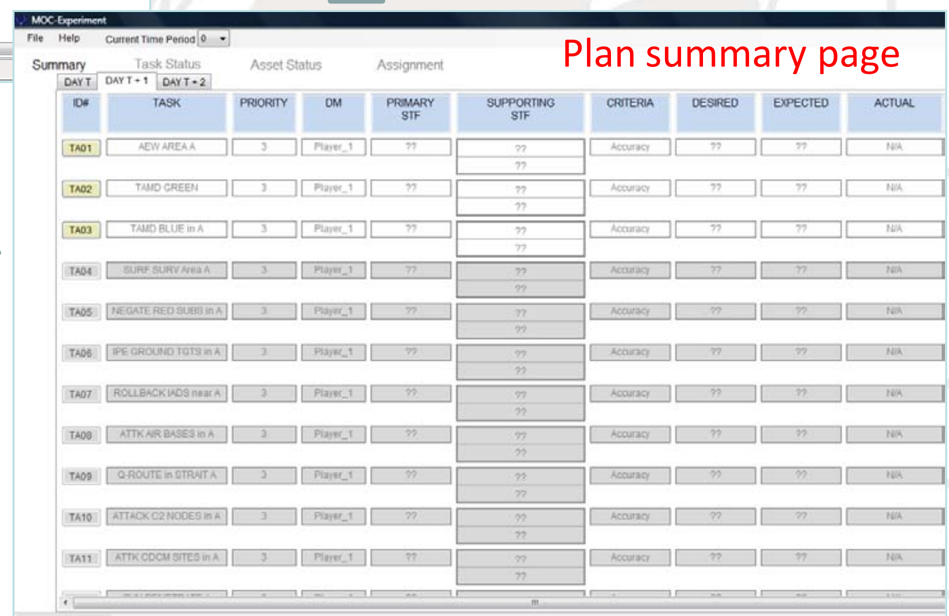
- FOPS assesses expected performance
- Modifies assignments on those tasks not meeting desired criteria
- When satisfactory, the plan is "finalized"
  - T+2 plan => start for next T+1 plan
  - T+1 plan => EXORD for tomorrow

Rolling Horizon Planning

For each task, the responsible FOPS planner assigns:

1) one primary TF
2) up to two supporting TFs
   - each in up to 2 warfare areas
3) desired perf level (accuracy, % complete)

The **plan** = aggregate of all assigned tasks for the given day, is posted on the summary



Experimental Design

Plan of execution

- The experiment occurs in five two-hour time blocks:
  - **Block 0** is an introduction to the experiment, including a brief on the mission, and initial training for players in their roles. The MOC Director (a confederate) leads the session.
  - **Block 1 is** the first full session where players are given the initial state of the scenario and are tasked with building plans for Block 2 (T+1) and Block 3 (T+2). These plans will be briefed to the MOC Director and submitted as a new plan at the end of this block.
  - **Block 2** begins with the implementation of the FOPS team plan produced in Block 1. The FOPS team will create a plan for Block 3 using their Block 1 plan for T+2 as a first guess. In this block, teams will also create the Block 4 (T+2) plan from the updated Block 3 plan.
  - **Block 3** implements the Block 2 plan for T+1, and teams use the Block 2 T+2 plan as the starting point for the T+1 (Block 4) plan. At the end of this session, another update briefing will be given to the MOC Director.
  - **Block 4** is the final session of the experiment. This session implements the Block 3 plan for T+1, and teams will use the Block 3 T+2 plan as the starting point for the T+1 (Block 5) plan. Expected progress will guide the T+2 (Block 6) plan, though neither the T+1 nor the T+2 plan will be executed.

Experimental Design

Independent Variables

		Information Content (a,b)	
		Forecast Only	Forecast with Explicit Uncertainty
Information Structure (I,II)	Automated Products	**Level Ia**	**Level Ib**
	Automated Products with Human Expertise	**Level IIa**	**Level IIb**

The levels of the independent variable as originally conceived separated content and structure of the forecast products into a classic 2X2 design.

Levels Ia and IIa represent much of the current practice in the Navy, Air Force and National Weather Service. Operationally, most atmosphere and ocean products are presented as deterministic forecasts with implicit uncertainty (Ia). This uncertainty is often clarified by additional information from experienced human forecasters (IIa). The Navy and Air Force are both considering moving to product portfolios with ensemble products (Level Ib or IIb).