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.

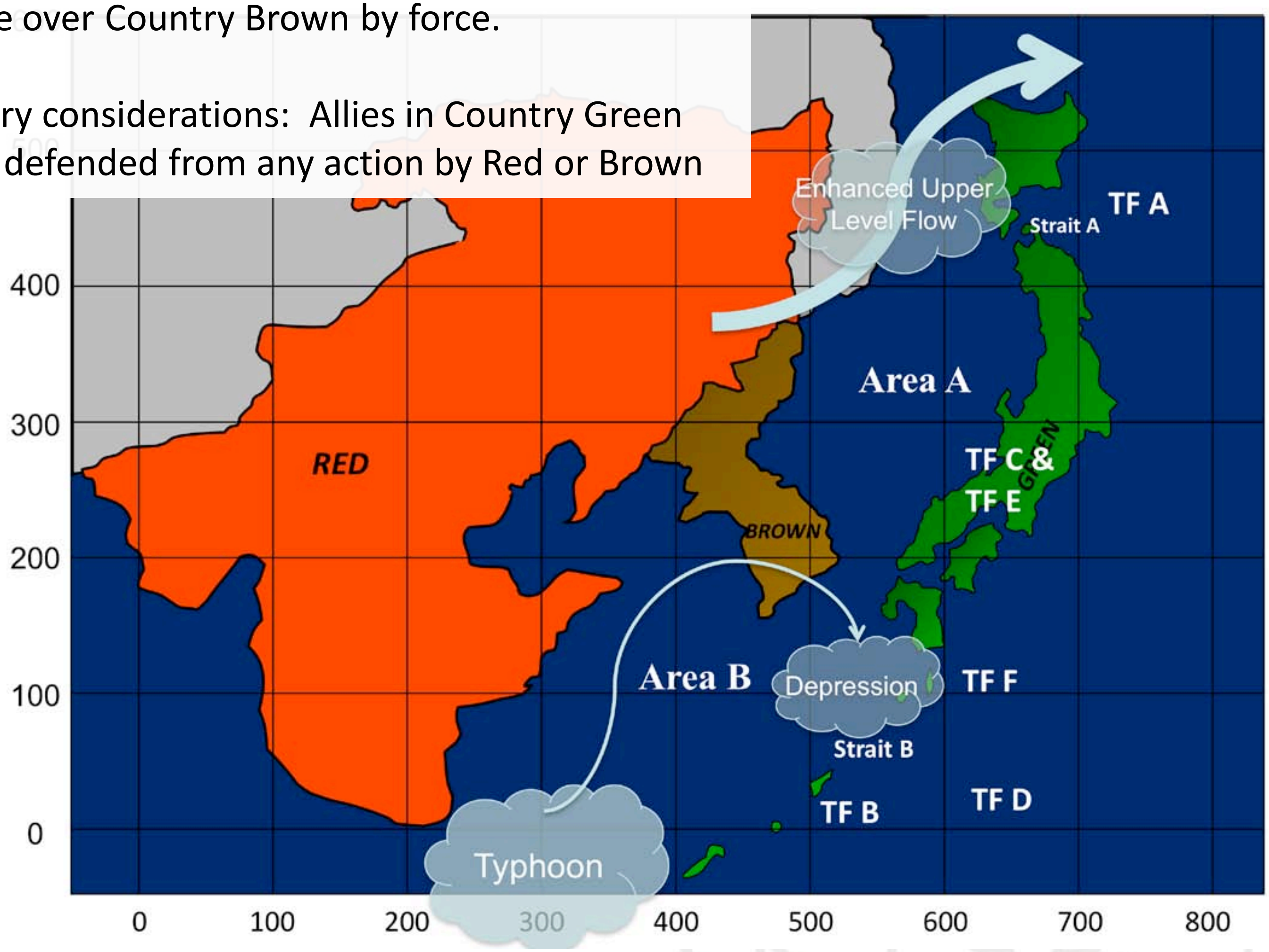
Research Questions

- 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.

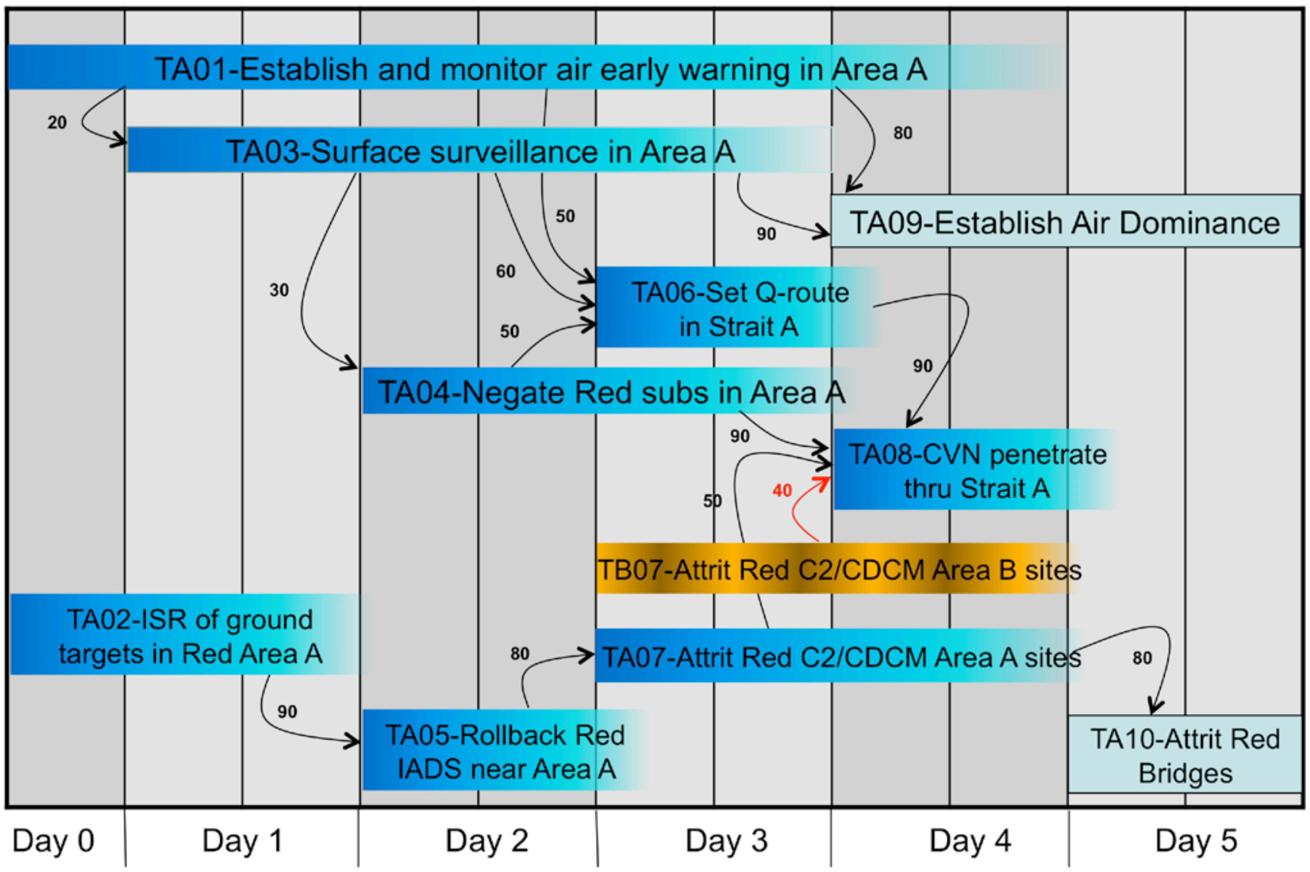
Secondary considerations: Allies in Country Green must be defended from any action by Red or Brown



Experimental Method

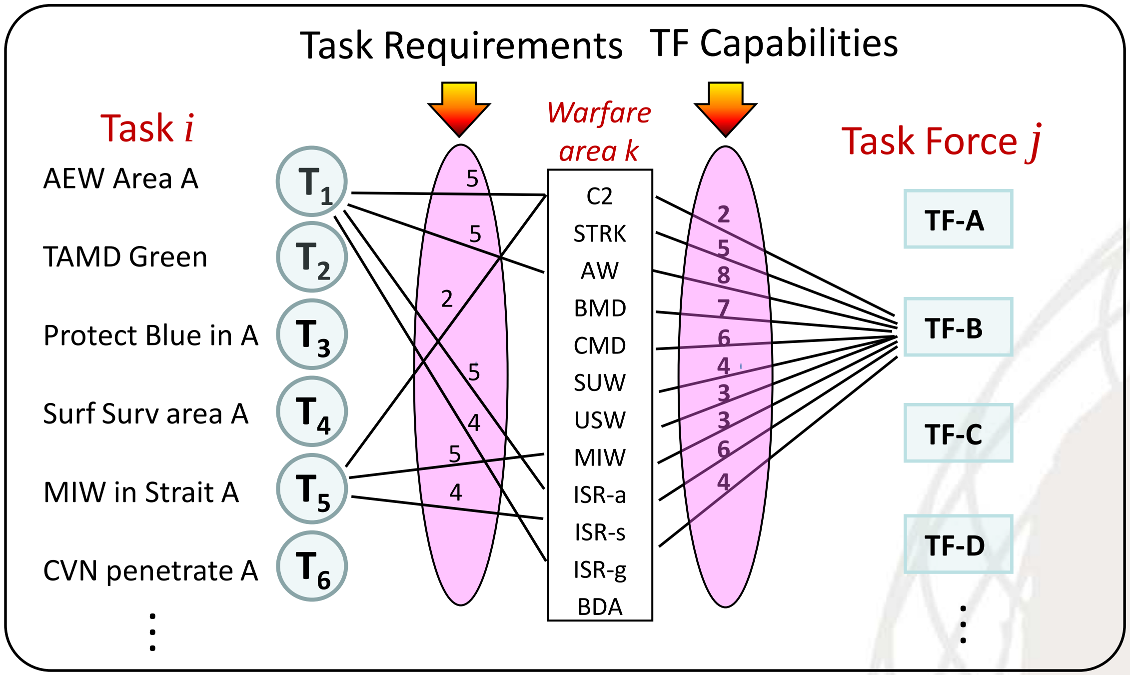
Modeling Formalism: Matching Assets to Tasks

Area A 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*

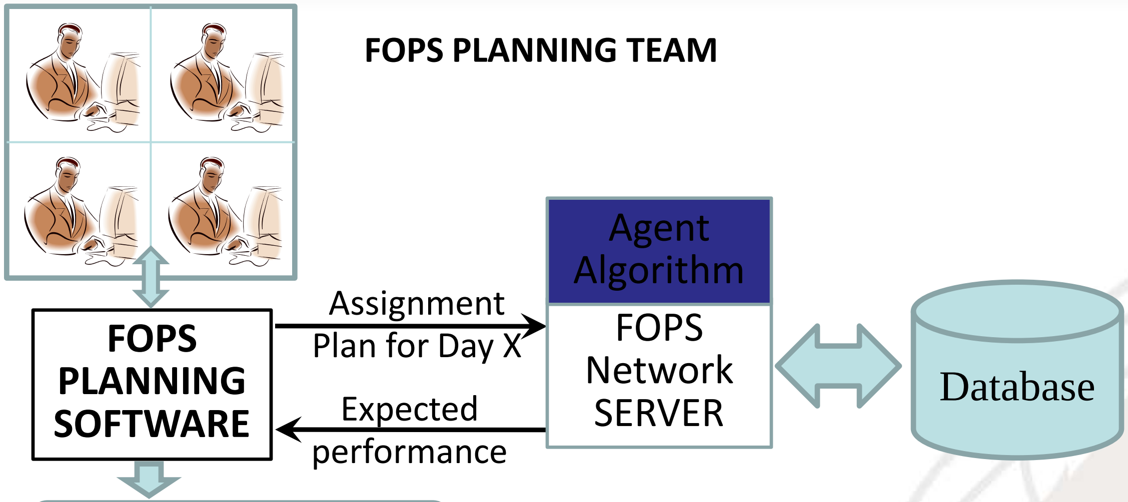
Values (resource vectors) are readily obtained via discussions with SMEs or Fleet personnel

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

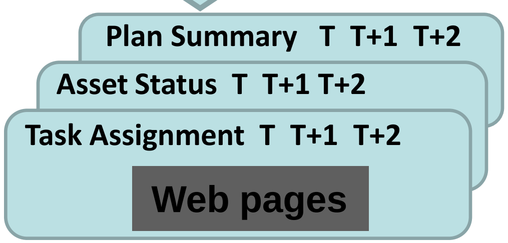
- Planner assignments (requests) include *performance goals* and *priorities*
- One task force is designated as primary by the FOPS planner
- Planner may assign other TFs as supporting in one or more warfare area
- The task forces (computer agents) determine how to best use assets to meet goals

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

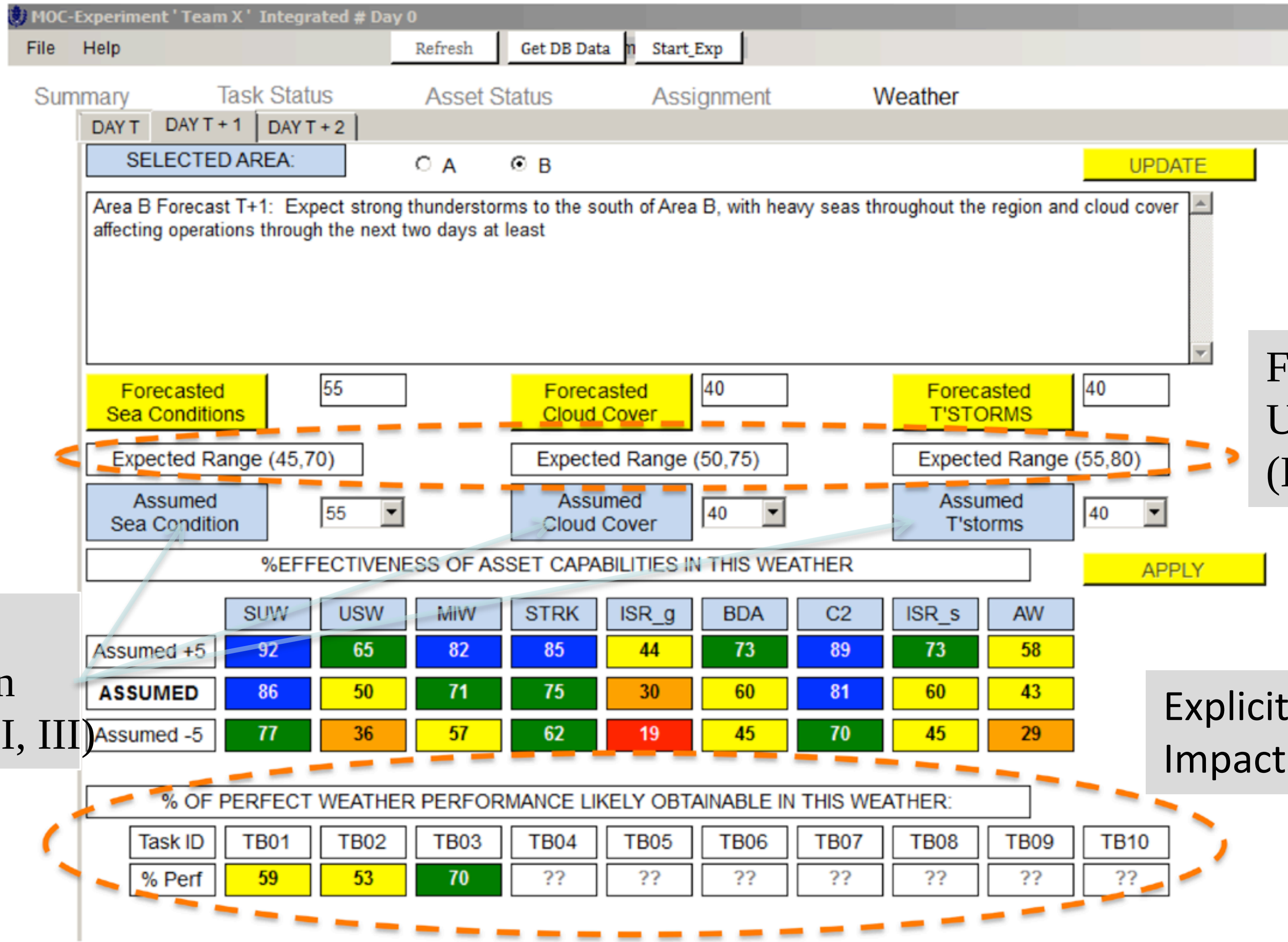
	Level I	Level II	Level III
Information Richness	*Forecast only, no additional weather or mission information*	*Forecast with explicit uncertainty information but no additional mission information*	*Forecast with explicit uncertainty and integrating mission information*

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

NAVAL POSTGRADUATE SCHOOL

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

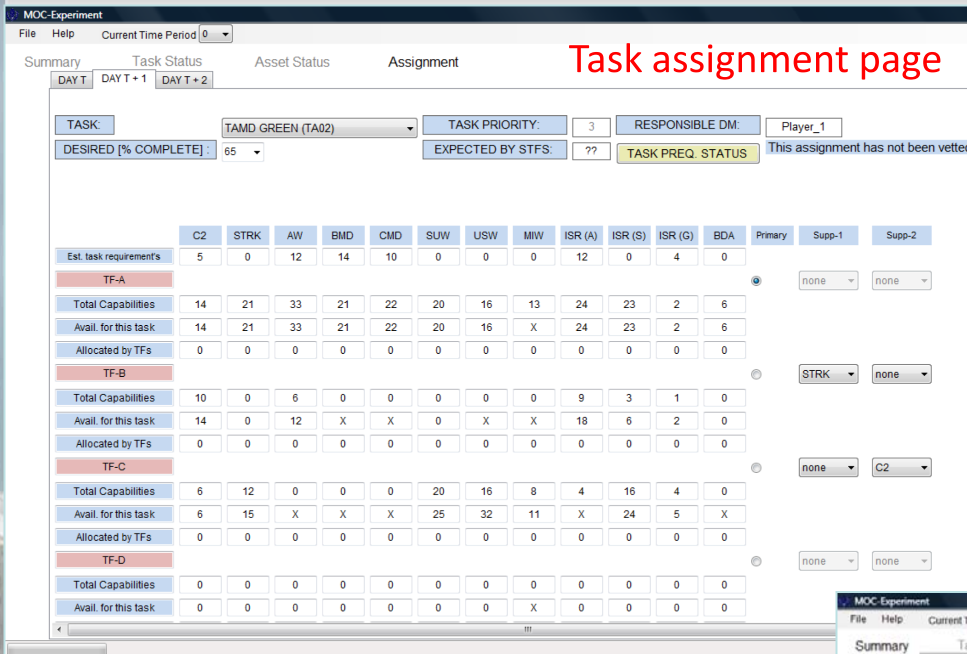
Backup Slides

Monterey, California
WWW.NPS.EDU

Experimental Design

Modeling Formalism: Matching Assets to Tasks

Task assignment page



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

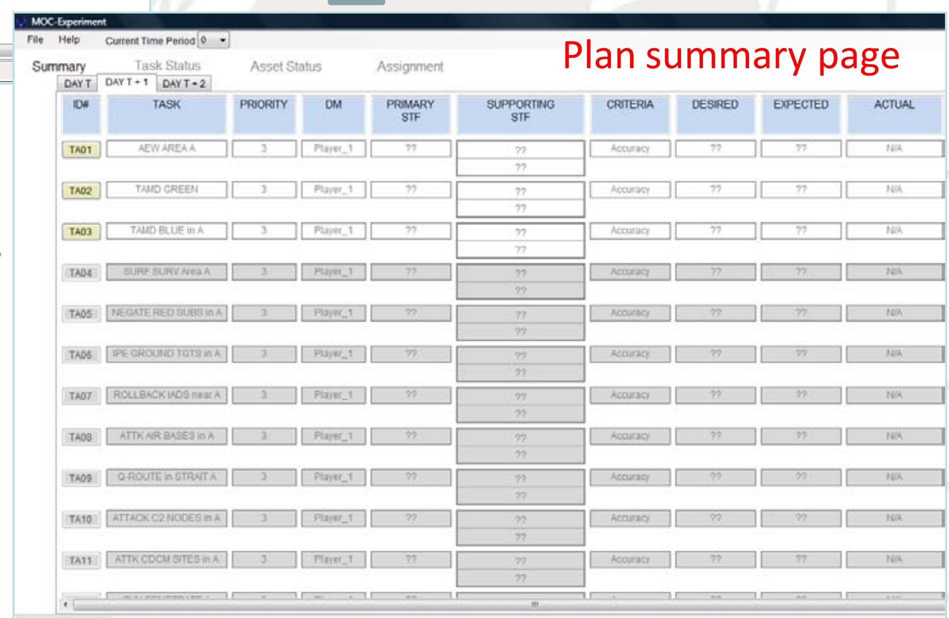
Rolling Horizon Planning

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

Plan summary page



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).