1996-06

Littoral Warfare Simulation Experiment

Kemple, William; Porter, Gary; Vodantis, Nick

Center for Advanced Concepts and Technology

Proceedings of the 1996 Command and Control Research and Technology Symposium: Command and Control in the Information Age, p. 536-543
https://hdl.handle.net/10945/64698

Proceedings of the 1996 Command and Control Research and Technology Symposium

(1996 : Monterey, Calif)

Command and Control in the Information Age

June 25 - 28, 1996
Naval Postgraduate School
Monterey, California

THE CENTER FOR ADVANCED CONCEPTS AND TECHNOLOGY

Littoral Warfare Simulation Experiment

William Kemple, Assistant Professor
Gary Porter Research Assistant Professor
Nick Vodantis, LT, USN and Graduate, C4I Systems Curriculum

Naval Postgraduate School
Monterey, CA 93943
408-656-3772
grporter@nps.navy.mil

Abstract

Fast Patrol Boats (FPBs) were considered a negligible threat when the U.S. Navy focused on blue water operations far from shore. They may pose a much greater threat now that the Navy's focus is in the littoral. Moreover, the threat may be greatly enhanced if the FPBs are able to employ Commercial Off-The-Shelf (COTS) Command and Control C2 equipment to coordinate their efforts.

This paper presents the design and results of a wargaming experiment conducted at the Naval Postgraduate School (NPS) to examine this issue.

The research question is how U.S. Navy Surface Action Groups (SAGs) operating in littoral waters would perform against FPBs equipped with COTS C2 equipment which would permit them to conduct coordinated saturation attacks. The experiment also looked at two other factors: SAG workload, which might exacerbate the effects of the FPB's C2; and the quality of intelligence our SAGs are provided which might help solve this increasing FPB threat.

A two-cubed factorial experiment was conducted to collect data on ten performance measures (PMs) for each of the 128 trials. Significant results were obtained for each of the three factors and three interactions over a broad range of PMs. These results are explained in operational terms.

1 Introduction

The details and results of a wargaming experiment conducted at the Naval Postgraduate School (NPS) during Winter 1996 are presented. The premise of the experiment is that the performance of a U.S. Navy Surface Action Group (SAG) conducting littoral warfare against fast patrol boats could be significantly affected if the enemy uses Commercial Off-The-Shelf (COTS) Command and Control (C2) equipment, similar to that used by the U.S. Navy to achieve the integrated C2 envisioned in C4I for the Warrior (C4IFTW).

Two other effects were also considered: the workload experienced by the SAG and the amount of intelligence that the SAG has available before hostilities commence.

A two-cubed factorial experiment was conducted to test three main hypotheses:

Ho1: Use of coordinated or uncoordinated tactics by red fast patrol boats makes no difference in blue SAG performance.

Ho2: Low or high SAG workload makes no difference in blue SAG performance.

Ho3: Area or Sector specific intelligence provided to the SAG makes no difference in blue SAG performance.

Four secondary hypotheses were also tested for interactions

A Persian Gulf scenario was developed as the context for the experiment. Eight vignettes were developed, each representing a unique combination of the factor levels. Common to all vignettes are the blue and red Orders of Battle (OOBs).

Each of the 128 trials was scheduled for 38 minutes of real time play. The game was played at four times real time to keep players occupied and to increase sample size within the overall time constraints.

B&R was used to implement the scenario and to serve as the stimulus for the subjects. Each subject controlled the blue side (SAG ships and helicopters). A B&R script and automated force routines controlled the remaining forces (red/hostile and white/neutral sides).

Sixteen officer students from NPS were used as subjects for the experiment. They represented all four services and had diverse backgrounds within each service as well

Ten performance measures automatically generated by B&R are used to test the hypotheses. The data analysis plan includes use of frequency plots, box plots, MANOVA, univariate ANOVA, interaction plots, residual plots, and non-parametric statistics to assess the data. A significance level of $\alpha = 0.05$ was established to test all null hypotheses.

Hypotheses tests show significant results in the main hypotheses and several interaction hypotheses over a broad range of performance measures.

2 Purpose – To Examine a Real World Problem

The U.S. Navy's missions are changing from blue water strategies that necessitated sophisticated and expensive multi-purpose combatants operating in carrier battle groups (CVBGs) toward missions requiring new regional and littoral strategies that frequently employ small groups of ships called Surface Action Groups (SAGs) to respond to these missions [Hughes, 1986]

The Fast Patrol Boat (FPB) has always been a cost effective platform for littoral warfare, especially for third world nations. In the past, potentially hostile FPBs have been viewed by the U.S. Navy as more of a nuisance factor that stresses C2 than a significant threat to the CVBG or SAG operations which conduct many missions without support of a CVBG's air wing. This is due in part to the FPB's poor C2 and lack of ability to conduct coordinated operations in the past.

Today's advanced COTS C2 technology could perhaps drastically increase the FPB threat. With effective C2 these third-world nation's FPBs could mount coordinated attacks that could saturate the SAG's defenses.

The FPB's small size and high speed combine to complicate U.S. intelligence gathering. In addition, U.S. Naval combatants are not optimized to track and target the FPBs, nor can they still assume the dominant tactical advantage inherent with superior C2. Still employing forces specifically designed for blue water warfare, the U.S. Navy is caught in the middle of this shifting paradigm and appears slow to react to the fast patrol boat (FPB) threat in the littorals.

3 Research Questions

FPBs already pose speed and stealth advantages over most naval combatants. The main research question is the effect that coordinated saturation attacks by FPBs, employing advanced C2, will have on an engagement with U. S. SAGs operating in the littoral seas. Also of interest are the effects of increased workload and incomplete intelligence on U.S. operations, both of which accompany the presence of FPBs in an operating area, and their interactions with the main effect.

4 Approach

A three factor experiment was constructed involving human decision makers. All factors were presented at two levels, high and low, resulting in a two cubed factorial experiment. The experiment was designed and executed on the *Battle Management Assessment System and Raid Originator Bogie Ingress* wargame, more commonly called Batman and Robin (B&R) [Federico, *et al.*, 1991]. Eight different scenarios were generated on Robin, the scenario generation capability of B&R, to present all combinations of levels and factors. Each scenario was then administered to NPS students on Batman, the executable portion of B&R, for a total of 128 trials. The geographic location of each scenario was identical: the Persian Gulf, between the Straits of Hormuz and Bahrain. Red forces were comprised of nine FPBs per scenario. A three combatant U.S. Navy SAG, including five organic helicopters, constituted Blue forces for each scenario. The three design variables or factors were C2 Coordination, Wokload, and Information Completeness.

4.1 FPB C2

The level of coordination or C2 used by the FPBs during an attack on the SAG is the primary factor and central area of interest in this experiment and the original impetus for the experiment. The analysis, presented later will focus on this factor. Two levels of FPBs C2 were achieved by varying their tactical coordination between high and low. High coordination was modeled by three concurrent waves of attacks by the FPBs. For this case seven of the nine FPBs were in restricted emission control (EMCON) status with the remaining two in an unrestricted EMCON status. It was assumed that the coordinated attackers would share information about the enemy, the U.S. SAG in this case, and thus require fewer emissions. Low coordination was modeled by sequential attacks of FPBs with only two of the nine FPBs in an EMCON status and the remaining seven emitting.

The two other factors in the experiment are workload and information completeness. While each was analyzed, the primary focus of the analysis will be on their interaction with the main effect, C2.

4.2 Workload

It is conceivable that third-world countries employing FPBs with COTS C2 could also use their advanced C2 to influence the level of non-hostile or neutral contacts within Blue's battlespace. For example a third world nation could orchestrate a massing of own or allied country non-hostile shipping near the SAG in conjunction with an FPB attack. This increased workload on Blue could have a significant impact on SAG combat performance. High workload scenarios contained 24 neutral ships in the operating area. These are in addition to the nine hostile FPBs. Low workload scenarios contained nine neutral ships in addition to the nine hostile FPBs.

4.3 Information Completeness

One way to counter the advantage gained by the FPBs use of COTS C2 is to increase the U.S. intelligence collecting capabilities. either through acquisition of new systems, or more likely, through repriortization of existing U.S. Navy and national intelligence assets. Two levels of intelligence, called information completeness, were used in the experiment, by providing the subjects with a maritime intelligence message at the beginning of each scenario. This message contained a high or low level of information completeness. High information completeness was represented by providing subjects the ratio of hostile ships in each threat sector to the total number of ships in the threat sector. For low information completeness, only the total number of ships per threat sector was given.

5 Anticipated Results

The hypothesized results were that the Blue SAG would perform better when the following occurred:

Incoming FPB attacks were uncoordinated (sequential) vice coordinated (concurrent); the ship's combat team was experiencing low workload conditions; and information or intelligence reports were more complete.

6 Experimental Design

The complexity of the scenarios for this experiment was limited due to time constraints and the varied backgrounds and limited operational experience of the subjects. Blue Force resources were initialized in the same manner for every scenario and for every subject. Otherwise, there was risk that subjects with operational experience would rely on their expert knowledge to position their forces in an advantageous manner based on platform capabilities. In addition, certain artificialities were accepted in the interest of practicality and usability. For instance, the wargame was run at a speed of 4:1 (i.e., 4 seconds of real time = 1 second of simulation time). Therefore a standard 152 minute scenario was completed in 38 minutes of real time. All scenarios were similar in construction except to allow for the systematic introduction of all levels of all factors. Scenarios were constructed to appear different to the subjects (without affecting essential traits). This diminished the chance of subjects relying on prior scenario knowledge to "game" the scenario instead of making decisions based on current scenario developments.

6.1 Physical Setup

B&R was obtained from the creator of the wargame, Dr. Pat-Anthony Federico at the Naval Personnel Research and Development Center, San Diego. It is written in "C" programming language. B&R was installed onto five Sun Sparc-20 stations in the NPS' Systems Technology Lab. All scenarios were created in Robin on one Sparc station and ported to the other four to ensure standardization. The subjects executed each scenario independently of the other subjects. The experiment was conducted in a controlled environment.

6.2 Subjects

Sixteen subjects participated in the experiment. This section discusses demographic information on the subjects and how the subjects were trained for and participated in the trials

6.2.1 Demographic Information

Of the 16 subjects, all were student officers assigned to NPS. Eight were USN, five were USMC, two were USAF, and one was USA. Six of the eight naval officers were designated Surface Warfare Officers (SWOs) and were thought to have the highest degree of expert domain knowledge based on the scenarios. Years of service ranged from 5-16 years.

A demographic questionnaire was also completed by each subject during the overview brief. The subjects were asked to subjectively rate their level of proficiency in each of the following areas on a scale from one (poor) to ten (expert). Detailed analysis of the demographic data was not attempted nor was an attempt made to evaluate their impact on the hypotheses or interactions. This is an area for further work. The demographic data elements are described below.

6.2.1.1 Naval Tactical Data System (NTDS) Symbology/Nomenclature

Proficiency with NTDS symbology was asked because the symbology used in B&R is very similar to NTDS symbology. It was presumed that familiarity with similar symbology would give a subject an advantage in adjusting to the wargame. As expected, the range of responses to this question were bimodal with ten of the sixteen subjects evenly split between the top end of the scale and the bottom end with naval officers dominating the high mode.

6.2.1.2 Simulations and Wargames

Experience with wargames could influence a subject's performance. The responses to the question regarding wargame experience was normally distributed.

6.2.1.3 Tactical Action Officer (TAO)

The TAO plays a key role in defending the SAG against an FPB attack and launching the counter-attack. Fifty percent of the subjects had no TAO experience.

6.2.1.4 Computer Proficiency

Finally, the data pertaining to the level of computer proficiency was collected.

6.2.2 Subject Training and Trials

All subjects were given a 50 minute overview brief of SAG tactics, platforms, sensors, weapons performance parameters, weapon status, weapon posture meanings, threats, and B&R rules. Each subject was given two practice trials followed by the eight experimental trials in which data was collected. During the practice trials, subjects received one-on-one instruction on playing B&R from trained proctors. All trials were completed during a ten day period. Subjects were limited to three trials per day.

6.3 Statistical Design of the Experiment

The experiment was counter-balanced to negate the learning effect. The two-cubed was implemented using eight unique scenarios. Using 16 subjects and a complete randomized block design with 2 replications yielded an 8x16 output matrix. To construct the design matrix the scenario type, order of scenario presentation, and subject's assignment to a scenario presentation were all randomized.

7. Data Collection Plan

This section discusses the performance measures that were automatically generated by B&R and used to test the hypotheses, and how the data were prepared for statistical analysis.

7.1 Measures

B&R's automatic data collection feature collected 81 performance measures (PMs) or dependent variables for each trial. The 81 PMs are divided into two broad categories: Blue Forces' performance vs. Red Forces and Red Forces' performance vs. Blue Forces. Each of these two broad categories is further divided into 3 sub-categories: air forces vs. surface, surface forces vs. surface, and combined air/surface forces vs. surface. Only 10 PMs were applicable to this experiment. The first five PMs measure Blue Forces' performance. The second five measure Red Forces' performance. The PM's used are defined below. The first five measure Blue's performance against Red. The second five measure Red's performance against Blue.

1. *BAFP 1: FPBs detected by Blue Air (%).*

2. *BAFP 2: Ave. range (measured from nearest Blue force ship) at which FPBs were first detected by Blue Air (nm).*

3. *BSP 3: FPB's destroyed by Blue Surface (%).*

4. *BSP 4: Ave. range (measured from nearest Blue force ship) at which FPBs were destroyed by Blue Surface (nm.).*

5. *BAFP 12: Number of TAO weapon assignments (count).**

6. *A 1: Blue Air detected by the FPBs (%).*

7. *S 1: Blue Surface detected by the FPBs (%).*

8. *A 2: Average range (measured from nearest FPB) at which Blue Air was first detected by the FPBs (nm.).*

9. *S 2: Average range (measured from nearest FPB) at which Blue Surface was first detected by the FPBs (nm.).*

10. *A 3: Blue Air destroyed by the FPBs (%).*

7.2 Data Preparation

B&R automatically calculates elemental values and statistics for the PMs. However hardware and software limitations prohibited the automatic transfer of information to a data file for further analysis. Therefore the data files were manually constructed. At the end of each trial B&R displayed the PMs on the display screen. Each subject transferred their performance data set from the monitor to the PM data collection sheet. The data were then manually entered into a spreadsheet.

Before any manipulations were performed,

* Blue forces had to utilize the TAO function in order to fire their weapons and engage Red forces.

the data were put through a rigorous quality control process to check for input errors. Next, data reduction was performed and the data ported from spreadsheet into a statistical package (Minitab) for statistical analysis. The data were then re-checked for errors.

7.3 Data Coding Scheme

For post data processing purposes, each scenario was uniquely described by the level, high or low, for each of the three factors: Information Completeness (I), Work Load (L), and FPB C2 (C). This information was coded as a three digit number with each digit limited to a '2' (high) or a '1' (low). All the data entered from each PM data collection sheet into the spreadsheet was indexed by this three digit scenario code which identified one of the eight scenario types from which it was derived. For example, the scenario with I, L, and C at the high level is coded 222. The scenario with all factors at the low level is coded 111. We anticipated scenario 211 (high information, low work load, low FPB coordination) would produce the best results for Blue forces

.

8 Analysis Plan

The data analysis plan includes frequency plots to obtain a pictorial representation of the 128 data points per PM (16 subjects x 8 scenarios). This established an initial feel for the data and provided insight for further detailed analysis. The shape of some of these plots made our initial assumption of normality for these data questionable. Next, box plots were used to examine the three factors' individual impacts on the PMs. After a preliminary understanding of the data were obtained, we tested the effects on all PMs simultaneously using three factor MANOVA. Since residual analysis showed that not all PMs appeared normal, we used parametric and non-parametric ANOVA and interaction plots to further assess the data. A significance level of $\alpha = 0.05$ was used to test all null hypotheses.

9 Hypothesis Results

All three factors and three of their interactions produced significant effects, although not for all PMs. All significant effects are described below.

9.1 Main Effects

All three factors produced significant effects and are described below with their associated PMs.

9.1.1 FPB C2

FPB C2 had a significant impact on seven PMs:

- *The percentage of FPBs detected by Blue Air (BAFP 1), p = .000, 90.1% detected when FPBs were coordinated, and 98.1% detected when FPBs were uncoordinated;*

- *The average range of detection of the FPBs (BAFP 2), p = .000, 76.0 nm when FPBs were coordinated, and 66.6 nm when FPBs were uncoordinated;*

- *The percentage of FPBs destroyed by Blue Surface (BSP3), p = .000, 65.5% destroyed when FPBs were coordinated, and 84.9% destroyed when FPBs were uncoordinated;*

- *The average range at which the FPBs were destroyed by Blue Surface (BSP 4), p = .002,32.8 nm when FPBs were coordinated, and 27.5 nm when FPBs were uncoordinated;*

- *The number of TAO weapon assignments (BAFP), p = .000, 17.0 assignments when FPBs were coordinated, and 22.3 assignments when FPBs were uncoordinated;*

- *The percentage of Blue Air detected by the FPBs (A 1), p = .012, 93.9% detected when FPBs were coordinated, and 97.9% detected when FPBs were uncoordinated; and*

- *The percentage of Blue Surface detected by the FPBs (S 1), p = .003, 19.3% detected when FPBs were coordinated, and 27.1% detected when FPBs were uncoordinated.*

When the FPBs had high C2, fewer were detected, however, of those detected the range of detection was greater. Similarly, when the FPBs had high C2 fewer were destroyed, but those that were destroyed were done so at a greater range. It only makes sense that if they were not detected they could not be killed, and if they were detected further out, that they would be killed further out.

But why would fewer be detected and those detected be detected at greater range? The answer to this is inherent in the scenario design and is two fold. When the FPBs had high C2, they conducted a coordinated (concurrent waves vice sequential) attack with seven of the nine FPBs in EMCON. Without the FPBs radiating they were more difficult to detect, but due to their organization Blue had more time to prepare and to deploy forces.

Fewer TAO assignments were made when the FPBs had high C2. This explains why fewer FPBs were destroyed when C2 was high. Analyzing the operations from the FPBs perspective reveals that the FPBs detected fewer Blue Air and Surface platforms when they had high C2. This is a direct result and a potential disadvantage of the high C2 (stealth) attack.

9.1.2 Workload

Workload had a significant impact on two PMs:

- *The average range of detection (BAFP 2), p =.001, 69.6 nm when blue workload was high, and 73.0 nm when blue workload was low; and*

- *The average range at which the FPBs were destroyed (BSP 4), p = .001, 27.4 nm when blue workload was high, and 32.9 nm when blue workload was low.*

In both cases, performance was adversely affected (range decreases) when workload increased.

9.1.3 Information Completeness

Information completeness had a significant impact on only on PM:

- *The average range at which the FPBs were destroyed by Blue Surface (BSP 4), p = .047, 31.8 nm when information was complete, and 28.5 nm when information was incomplete.*

With complete information Blue Forces were able to kill the enemy over 3 nm further out than with incomplete information.

9.2 Interactions

Three interactions produced significant effects and are described below.

9.2.1 Information Completeness and FPB C2

The combined effects of Information Completeness and FPB C2 had a significant impact on the mean range at which the FPBs were destroyed by Blue Surface (BSP 4), p = .011.

When FPBs were conducting uncoordinated attacks (low C2), information completeness has no effect. But when the FPBs use coordinated attacks (high C2), the FPBs are destroyed at much greater range when Blue has complete information, 36.5 nm, than when Blue has incomplete information, 29.0 nm.

9.2.2 Workload and FPB C2

The combined effects of Work Load and FPB C2 had a significant impact on the average range at which Blue Air was first detected (A 2), even though neither taken alone is significant. Additional analysis does not yet help explain why the FPBs would have detected Blue Air further out with high C2: p = .000, 46.14 nm when C2 high, and 44.56 nm when C2 low.

9.2.3 Information Completeness, Work Load, and FPB C2

The combined effects of Information Completeness, Work Load, and FPB C2 had a significant impact on the average range of detection of the FPBs (BAFP 2) although no attempt is made in this preliminary analysis to interpret this interaction.

9.3 Operational Interpretation

Five Blue helicopters scouting a geographically constrained littoral region probably reduced the importance of judicious resource allocation and thus lessened the effect of whether Blue had Complete or Incomplete Information. In essence, the subjects had nearly total radar coverage provided by the helicopters regardless of where they directed them.

The impact of FPB C2 changing levels registered in more dependent variables than any other factor. Probable causes include: 1.) FPBs were in EMCON when C2 was high which greatly effected there ability to *see* and be *seen*. and 2.) When FPB C2 was high their tactics differed significantly than when their C2 was low. In the former case, the FPBs converged on Blue in three concentric circles. The attack was concurrent in three waves. In the latter case, the attack was more sequential and random. The net result might be that even though Blue was able to carry out a more methodical and deliberate counterattack when the FPBs had high C2, they could not detect as many and therefore could not kill as many.

10 Recommendations

The experiment could be improved by more realistically setting the levels of information completeness and workload. Reducing the number of helicopters would help accomplish the former for reasons stated above. Workload could be varied between high and low based on real world maximum and minimum shipping levels in the Persian Gulf to ensure applicability to present day operations. The FPB threat is an emerging one, so it will be more difficult to find real world data on which to model FPB C2 and tactics, but this aspect of the study should be pursued.

The easiest way to extend this analysis is to segregate subjects by warfare specialty or service and then analyze the data accordingly. Other ways include: replicating the experiment on another simulation system to show the affects of B&R artifacts, using different scenarios to neutralize scenario effects, or investigating different factors or different levels of studied factors.

The analysis of the data collected from this experiment is ongoing. It is possible that different conclusions may be reached when more detailed analysis is completed.

References

[Hughes, 1986] Wayne P. Hughes, Jr., USN (Ret.), *Fleet Tactics*, Naval Institute Press, Anapolis, Maryland, 1986.

[Federico, *et al.*, 1991] Pat-Anthony Federico, Randy R. Ullrich, Brian L. Van de Wetering, Chistel L. Tomlinson, Dean J. E. Long, Fred R. E. Long, and Thomas E Bridges. *Human-Computer Interfaces for Tactical Decision Making, Analysis, and Assessment Using Artificially Intelligent Platforms.* NPRDC, San Diego, California, 1991.