12th ICCRTS

"Adapting C2 to the 21st Century"

Measuring Team Collaboration in a Distributed Coalition Network

Network-Centric Experimentation and Application Track

Elizabeth K. Bowman
Army Research Laboratory
Building 459
Aberdeen Proving Ground, MD 21005
410-278-5920
ebowman@arl.army.mil

Measuring Team Collaboration in a Distributed Coalition Network

Abstract

Distributed teams representing multidisciplinary perspectives and operating in a collaborative information environment will define the future of Command and Control (C2). Multinational Experiment 4 (MNE 4) provided researchers an opportunity to evaluate how distributed teams interact in a collaborative, networked environment to conduct the Effects Based Approach to Operations (EBAO). Several factors related to team collaboration were evaluated. These included measurements of workload, perceptions of information quality, ability of subjects to develop trust in teams and the staff process, team process dynamics and the assignment of roles and responsibilities within teams.

Results show that distributed teams were able to establish trust, but not all teams were equally effective in assigning roles and responsibilities and in establishing inter-team relationships. This resulted in confusion about the overall objective. Teams also reported an unequal distribution of effort, performance, and frustration. Larger teams reported increased workload, lower performance, and higher frustration. Respondents' ratings of information quality suggest that information was complete and sufficient, but less timely, understandable, and accurate. These findings are summarized and recommendations made for organizational design, collaborative system use, and training.

Introduction

The response to asymmetric warfare demands military collaboration with allied nations, other-government agencies and non-government humanitarian organizations. Advanced networking technologies provide the means to achieve this collaboration in a distributed manner. One process these agencies may use to coordinate this response is the Effects Based Approach to Operations (EBAO). In EBAO, operations are focused on influencing or changing system behavior or capabilities using selected instruments of power in order to achieve directed policy aims (for a more detailed description, see the MNE 4 EBO Concept of Operations, US JFCOM, 2004). A central focus of the EBAO is to achieve a greater unity of effort among the various instruments of power. In a coalition headquarters, this translates into the need for effective teamwork in a multidisciplinary staff.

The EBAO is fundamentally different from traditional military tactics of attrition in that it integrates all domains of national power (political, military, economic, social, information, infrastructure) to influence the behavior of other actors (Smith, 2002; US JFCOM, 2006). One predominant feature of this process is the integration of non-military actors into the traditional military staff group. The computing power of network architectures allows this staff to function in a distributed mode, making it possible for teams to form on an ad hoc basis. In this emerging era of team work, we examined how the distributed nature of communication would impact team collaboration.

Multinational Experiment 4 (MNE 4) was conducted in February 2006 with the cooperation of many nations and the North Atlantic Treaty Organization (NATO). The experiment included two separate headquarters conducting EBAO; one was the U.S.-based multinational Coalition Task Force (CTF) and the other was the NATO Response Force (NRF). The CTF was organized in a distributed configuration and the NRF was arranged in a co-located manner. This report describes the CTF results only.

In the CTF, partner nations identified participants to play military and civilian staff roles in a simulated combined joint task force. Most participants were physically located in their respective nations and connected to others via a collaborative network. [1] Each participant used the collaborative system to interact with their distributed team members.

During experiment execution, each team in the U.S.-based headquarters group was assigned a space where they could be co-located. The picture in Figure 1 depicts a small portion of the entire experiment space. Group members had contiguous seating assignments but conducted most business over the collaborative system. As explained earlier, the U.S.-based group comprised the multinational leaders of the staff groups and the U.S. staff. For example, the Effects Based Execution group was led by a U.K. officer and included 21 members. The U.S. portion of this group was 4. During execution, the EBE group size at the U.S. location was 5 (the four U.S. staff and the U.K. lead). The remainder of the EBE group was located in their respective nations, and conducted business over the collaborative software system. [2]

Figure 1 The CTF experiment layout.

The CTF was comprised of several teams. These included the Command Group, Effects Based Planning (EBP), Effects Based Execution (EBE), Effects Based Assessment (EBA), Red/Green teams (R/G), Knowledge Superiority (KS), Multinational Interagency Group (MNIG) and System of System Analysts (SOSA). The Red and Green Teams were sub-groups of the EBP group and the SOSA group was a subset of the KS group. They were broken out for this analysis

[1] Participants from the other nations who were filling staff leadership roles were co-located in the U.S. with the command staff.

[2] The two exceptions to this were the Swedish and Finnish delegations. Due to technical constraints, these delegations were located in remote rooms in the JFCOM building but were connected to the other participants via the collaborative system.

because they functioned as separate groups due to the nature of their responsibilities. A brief description of these teams follows.

The CG, EBP, EBE, and EBA groups had responsibilities similar to current military staffs today. These groups were responsible for command, planning, execution, and assessment of current operations, respectively. The MNIG was an interagency group comprised of coalition nations' representatives from the diplomatic, economic, social, and information domains. The Red and Green teams were small sub-groups of the planning staff and were responsible for developing responses to the CTF plan from the enemy and friendly perspectives.[3] The KS group included the Knowledge Base Development (KBD) and Knowledge Management (KM) staffs. The KBD group was responsible for maintaining current information about the area of operations. The KM staff was responsible for maintaining the database structure supporting the push and pull of information from the knowledge base. The SOSA group was organized along the PMESII spectrum. [4] At least one SOSA was assigned for each of the PMESII domains and contributed to the KBD function. During the experiment, the SOSAs were responsible for providing their specific analysis expertise as requested by members of the staff.

In our examination of team cohesion in distributed and *ad hoc* settings we focused on the ability of the staff teams to form a trusting relationship (Fine & Holyfield, 1996; Mayer, Davis & Schoorman, 1985) and the ability of team members to work together and to share knowledge of individual and group role sets (Warne et al., 2004). We were also interested in documenting the expected large amounts of information that can overwhelm a staff working in a networked environment (Edmunds & Morris, 2000).

Method

In this experiment, we chose to focus on the following human factors issues to examine team cohesion in EBAO. These were balanced roles and responsibilities, quality of information, trust in process, organization, and technology, workload, and teamwork. We hypothesized that if high levels of these variables could be achieved, the coalition staff would be able to approximate a cohesive staff. We defined a cohesive staff as one capable of meeting deadlines by dividing responsibilities and of handling tasks in an efficient way. We also used a demographic survey to collect information on participants. This survey was completed at the beginning of the experiment.

The NASA Task Load Index (NASA TLX) was used to measure workload because it is easy to administer, is accepted by respondents, and has high face validity (Hill, Iavecchia, Byers, Bittner, Zaklad, and Christ, 1992). On four days of each of the three weeks in the experiment, participants were asked to complete this survey at the end of the day. The survey asked respondents to rate their perceived workload experienced that day in terms of mental and physical workload, time pressure felt, satisfaction with own performance, effort, and frustration felt. Each of these questions is answered on a 10 point scale (1=low, 10=high).

[3] Red is a term referring to the enemy; green is a term referring to friendly entities.
[4] PMESII: Political, Military, Economic, Social, Information, Infrastructure

Five questions were developed to measure participants' perceptions on their ability to trust in their ad hoc teams and the technology in use. These questions were 1) my team was open to ideas from all, 2) I was comfortable sharing ideas with my team, 3) team members were kept informed, 4) the collaborative technology made it possible for my ideas to be understood, and 5) the collaborative technology was an efficient way to work in a distributed environment. Each question was answered on a 7 point scale (1=low, 7=high).

Ten questions were used to measure participants' perceptions on the quality of information they received in MNE 4. These questions were all measured on a 7 point scale (1=low, 7=high). The questions asked participants to rate the quality of information they received on the basis of accuracy, appropriateness, accessibility, relevance, timeliness, completeness, sufficiency, conciseness, interpretably, and understandability.

Four questions were used to measure participants' perceptions of the team process they experienced in MNE 4. These questions were all measured on a 7 point scale (1=low, 7=high). The questions were 1) my team was effective in sharing information, 2) my team was effective in assigning roles, 3) my team was effective in assigning responsibilities, and 4) my team was effective in communicating ideas. The questions in this survey regarding roles and responsibilities do not conflict with the survey described below because HF4 asks if the team engaged in the process of assigning roles and responsibilities, not if a participant actually understood their contributions to the team.

Four questions were used to measure participants' understanding of their contributions to their primary team, or their understanding of their roles and responsibilities and those of their teammates. The questions were all measured on a 7 point scale (1=low, 7=high). These questions were 1) In my team, I was clear what was expected of me, 2) In my team, I was clear about what others were to do, 3) In MNE 4, I was clear about what other groups were to do, and 4) In MNE 4, I was clear about how all groups should work together.

131 participants participated in this experiment. Of the 124 respondents who answered the demographic survey, 68 (55%) had previous experience working in a distributed collaborative environment and 88% reported military experience. 66% reported participation in one of the spiral events (e.g. workshops or limited objective experiments) leading up to this experiment.

The primary data collection methodology was computer-based survey administration. These were analyzed with inferential statistics. The workload data were analyzed with repeated measures Multivariate Analysis of Variance (MANOVA), while the others were analyzed with analysis of variance (ANOVA). In addition, the researchers had access to real-time voice and typed conversations occurring in the small groups over the collaborative interface. This access essentially allowed the researchers to monitor several groups at one time and to record comments relating to an area of interest. Additionally, the survey questions mentioned above allowed respondents to provide explanations for any quantitative answer.

Results

Demographics

The CTF participants included 131 staff players, 124 of whom responded to the demographics survey (a 95% response rate). As shown in Table 1, the largest group was the Knowledge Superiority (KS), followed by EBP, EBE, EBA, and MNIG. The MNE 4 event was preceded by a two-year period of planning meetings and limited objective experiments. Table 1 also provides player reports of their involvement in one or more of these lead-up events. Fully 66% of respondents indicated they had participated in one or more of the 13 events preceding MNE 4. At least half of each group reported experience with the MNE 4 execution campaign prior to this experiment.

Table 1 Role Assignments and Experience

	EBAO Group						Total
Group Assignment	KS	EBP	EBE	EBA	MNIG	Command Group	
	39	32	21	14	13	5	124
% of group participating in one or more lead up events	73%	66%	52%	50%	54%	100%	66%

The breakdown of nationalities represented in the MNE 4 CTF staff is shown in Table 2. The U.S. and German staffs were the largest in the coalition.

Table 2 Represented nationalities in CTF

		Affiliation							Total
		Canada	Finland	France	Germany	Sweden	United Kingdom	United States	
EBAO Group	KS	3	2	4	16	4	2	8	39
	EBP	4	2	4	4	2	3	13	32
	EBE	3	1	2	5	2	4	4	21
	EBA	2	1	1	2	2	1	5	14
	MNIG	1	1	3	2	1	1	4	13
	Command Group	2	0	1	0	0	0	2	5
Total		15	7	15	29	11	11	36	124

The participants reported a mixture of active duty military, civilian employee, and contractor status at the time of their involvement in the experiment. Table 3 shows a breakdown of these categories. The 'other' category included the MNIG members who had affiliations outside of their nations' defense establishment.

Table 3 Affiliation of CTF participants

		Current Work Status							Total
		Air Force	Army	Civilian Government Employee	Contractor	Marine Corps	Navy	Other (describe)	
EBAO Group	KS	8	11	4	9	1	4	2	39
	EBP	4	9	2	9	2	5	1	32
	EBE	4	8	1	4	2	1	1	21
	EBA	1	2	3	5	0	2	1	14
	MNIG	0	1	5	4	0	1	2	13
	Command Group	1	1	0	3	0	0	0	5
Total		18	32	15	34	5	13	7	124

Of the MNE 4 CTF participants, 86% reported that they had military experience. Not unexpectedly, this is lowest in the MNIG group.

Table 4 Reported military experience

		Military Experience		Total
		Yes	No	
EBAO Group	KS	34	5	39
	EBP	31	1	32
	EBE	20	1	21
	EBA	11	3	14
	MNIG	7	6	13
	Command Group	4	1	5
Total		107	17	124

Of those participants reporting military experience, the data were skewed toward the upper end of the spectrum, as shown in Table 5. 84% reported 16 or more years of military experience.

Table 5 Years of reported military experience

		Years of Military Experience							Total
		1 to 5 years	6 to 10 years	11 to 15 years	16 to 20 years	21 to 25 years	26 to 30 years	More than 30 years	
EBAO Group	KS	1	3	5	10	10	4	1	34
	EBP	0	1	0	7	15	4	4	31
	EBE	0	0	3	6	5	3	3	20
	EBA	0	1	0	0	4	6	0	11
	MNIG	2	1	0	1	0	3	0	7
	CG	0	0	0	0	1	0	3	4
Total		3	6	8	24	35	20	11	107

However, when asked how many of these experienced participants had *multinational* military experience (see Table 6), the trend is in the reverse direction, with 21% reporting no such experience, 22% reporting less than one year's experience, and 43% reporting 1-3 years or less experience.

Table 6 Years of reported multinational military experience

		Years of Multinational Military Experience						Total
		None	Less than 1 year	1 to 3 years	4 to 6 years	7 to 10 years	More than 10 years	
EBAO Group	KS	8	9	12	2	3	0	34
	EBP	5	7	16	2	1	0	31
	EBE	4	5	9	0	1	1	20
	EBA	4	1	5	1	0	0	11
	MNIG	1	2	2	2	0	0	7
	CG	1	0	2	1	0	0	4
Total		23	24	46	8	5	1	107

Workload

In the workload analysis, the independent measures were EBAO group, week, and day. Dependent measures were the six aspects of workload: mental and physical workload, time pressure felt, satisfaction with own performance, effort, and frustration felt. Reported workload levels in the CTF EBAO headquarters showed that effort and performance were unequally distributed among groups. A repeated measures MANOVA was used to analyze these data, collected from participants at the end of four days in each of three weeks of the experiment. The data for time were separated into *week* and *day*. This analysis has two within factors, *week* and *day*, and one between factor, EBAO *group*. The N for this analysis was 92 due to missing data. Participants who did not complete the survey on each of the 12 days were deleted from the

analysis. Researchers did not replace missing values in the data set. This analysis tested the main effects of *day*, *week* and *group*, three two way interactions, and one three way interaction. The MANOVA revealed several significant main effects and interactions for workload. The significant effects are:

- Group main effect: significant *Wilk's* λ F (42 ,374) = 1.67, p =.005
- Week main effect: significant *Wilk's* λ F (12,73) = 6.98, $p < .0005$
- Day main effect: significant *Wilk's* λ F (18,67) = 2.65, p = .002
- Week * group interaction: significant *Wilk's* λ F (84,455) = 1.41, p = .016
- Day * group interaction: non-significant *Wilk's* λ F (1216,450) = .98, p = .538
- Week * day interaction: significant *Wilk's* λ F (36,49) = 2.60, p = .001
- Week * day * group interaction: non-significant *Wilk's* λ F (252,350) = 1.17, p = .088

The results shown above indicate the following. The EBAO groups differed in their perceptions of experienced workload. Participants also rated their workload by week and by day differentially. The week by group interaction suggests that some groups rated their workload differently by week. The week by day interaction suggests that participants rated their workload differently by days in the week. However, the MANOVA indicates only that overall differences exist. Further univariate testing was needed to understand how the groups vary with respect to these main and interaction effects. In analyzing these results, it is only appropriate to solve for the interaction effects, since the main effects are contained in the interaction equation (Stevens, 1996).

To evaluate the week * group interaction, a univariate analysis of variance (ANOVA) was used to determine which component of workload contributed to the interaction. Results showed significant effects for the *satisfaction with own performance* (F (14,7) = 2.91, p = .001) and *frustration felt* (F (14,7) = 2.04, p = .018) dimensions of workload. To further analyze these factors, ANOVA was performed for the significant workload components as the dependent variables and *EBAO group* as the independent variable. The results are reported below.

To further assess the significant effect, post hoc tests were conducted by controlling for week and assessing group differences during that week. As shown below, ratings of satisfaction with own performance were significant for performance in weeks one, two, and three. Ratings for frustration felt were significant only in weeks one and two but not three.

- Wk1perf, significant F (7,96) = 3.72, p = .001
- Wk2perf, significant F (7,99) = 2.90, p = .008
- Wk3perf, significant F (7,96) = 2.26, p = .036
- Wk1frust, significant F (7,96) = 3.27, p = .004
- Wk2frust, significant F (7,99) = 1.89, p = .018
- Wk3frust, non-significant F (7,95) = 1.14, p = .343

Post hoc tests for the significant univariate t-tests showed the following comparisons for satisfaction with performance and frustration. These tests show how the EBO groups differed with respect to their ratings of these workload dimensions, identified by the mean rating for the group. These results are shown in table 1 and 2 below. For each table, the group identified in

column two ('reference group') had a significantly different mean score than the groups identified in column 4 ('comparison groups'). Groups not included in column four had scores that were not statistically different than the column 1 group.

Table 7 Significant differences between groups on Satisfaction with Performance

Satisfaction with Performance Dimension					
Week	Reference Group	Mean	Comparison Group	Mean	Level of Significance
1	Command Group	7.31	EBP	5.44	$p = .032$
			EBE	4.38	$p = .001$
			EBA	5.48	$p = .050$
			KM	5.40	$p = .034$
			SOSA	5.02	$p = .015$
	EBE	4.38	EBP	5.44	$p = .030$
			MNIG	6.28	$p = .006$
	Red/Green Team	7.41	EBP	5.44	$p = .008$
			EBE	4.38	$p = .000$
			EBA	5.48	$p = .017$
			KM	5.4	$p = .009$
			SOSA	5.19	$p = .003$
2	Command Group	7.94	EBE	4.86	$p = .001$
			MNIG	5.81	$p = .029$
			SOSA	5.27	$p = .007$
	EBE	4.86	EBP	6.19	$p = .008$
			KM	6.32	$p = .008$
			Red/Green	7.20	$p = .006$
	SOSA	5.27	Red/Green	7.20	$p = .033$
3	Command Group	8.13	EBP	6.16	$p = .023$
			EBE	5.28	$p = .002$
			MNIG	6.19	$p = .050$
			SOSA	5.64	$p = .009$
	EBE	5.28	EBA	6.75	$p = .020$
			KM	6.57	$p = .014$

The pairwise comparisons shown in table 1 show that in the three weeks of the experiment, there were many significant differences among the EBAO groups with respect to their subjective ratings of 'satisfaction with own performance'. The Command Group had consistently high ratings for their own performance. We believe that this may be an artifact of the relatively unchanged roles for the Command Group in EBAO as experienced in MNE 4. For the staff participants in MNE 4, most roles required very different process tasks for the completion of the operational goals. These differences were required not only in the types of activities required but also in the use of tools and technology to complete actions and format documents. This was not true of the command group. These leaders (Commander, Deputy Commander, and Chief of Staff) performed essentially the same roles that they would have been familiar with in current operational assignments.

Also of note, the EBE group rated their performance significantly lower than other groups in all three weeks. A review of observations documented by embedded analysts suggests that these ratings are the result of the difficulties the EBE group experienced with the experiment scenario and the lack of information from the white-celled component commanders. [5]

Table 2 shows subjective ratings of 'frustration felt' for EBAO groups. In week one, the EBP group felt significantly higher frustration than EBA and MNIG. Based on documented analyst observations, this is due to the emphasis on planning at the beginning of the experiment (other groups had less work to do at this early point). Also in week one, the MNIG reported significantly lower frustration than EBE, KM, Red/Green team, and SOSA. Documented observations suggest this reflects the slow start for the MNIG group.[6] In week two, the Command Group reported significantly lower frustration levels than EBP, EBE, KM, MNIG, and SOSA. During week three, EBA reported significantly lower frustration than EBP. Documented observations suggest that the lower frustration levels for EBA are related to the lower group size, the more focused tasks for this group, and the ability of the EBA Chief to provide guidance and direction to his staff while faced with uncertainty.

Table 8 Significant differences between groups on Frustration Felt

Frustration Dimension					
Week	Reference Group	Mean	Comparison Group	Mean	Level of Significance
1	EBP	6.72	EBA	2.35	$p = .028$
			MNIG	3.81	$p < .0005$
	MNIG	3.81	EBE	6.57	$p = .001$
			KM	5.61	$p = .027$
			Red/Green	6.79	$p = .004$
			SOSA	6.64	$p = .002$
2	Command Group	3.0	EBP	5.75	$p = .007$
			EBE	6.26	$p = .002$
			KM	5.08	$p < .047$
			MNIG	5.54	$p = .021$
			SOSA	5.89	$p = .01$
3	EBP	5.44	EBA	3.95	$p = .039$

Univariate ANOVAs were used to determine what dimension of workload contributed to the significant interaction of week * day. ANOVAs showed a significant effect for time of

[5] The EBE group documented their need for more tactical information from the Component Commanders (played by the white (experiment control) cell) and stated that they could not adequately perform their mission without this input.

[6] Because the MNIG was staffed by actual interagency representatives from the participating nations, it was difficult for these individuals to devote one month to the experiment. As an accommodation, the full MNIG play was scheduled for the third and fourth weeks of the event (week one was training and week two was the first week of actual staff work). It was not until the last two weeks of the experiment that the MNIG was fully staffed and challenged with EBAO tasks.

workload measurement on reported mental workload, F (6, 504) = 2.82, p = .01, reported physical workload, F (6, 504) = 5.04), p = .0005, reported time pressure felt, F (6, 504) = 8.17, p = .0005, and reported effort F (6, 504) = 7.66, p= .0005. Univariate tests showed that daily workload reports for the dimensions identified above were significantly different in week one and week three. Further paired t-tests showed that this difference was primarily due to significant ratings for each dimension between the first day of week one and the last day of week three. Reviews of documented observations explain this trend. Participants were eased slowly into and out of the pace of the experiment on the first and last days of each week.

TRUST IN TEAM AND TECHNOLOGY

The distributed, ad hoc teams created for the CTF Headquarters in MNE 4 reported high levels of trust in their primary team and in collaborative technology employed during the experiment. The Trust in Team ANOVA revealed no significant differences between EBAO groups. Table 9 displays the survey questions and the overall mean rating for each question. The questions were scored on a 7-point scale (1=low, 7=high). Anecdotal observations and review of documented observations suggest that these high reported levels of trust are related to the fact that 66% of participants knew each other prior to the start of the experiment. Without this prior experience, one could reasonably expect distributed teams to develop trust over a longer period of time and with less unanimity.

Table 9 CTF Trust in Teams Responses

Survey Question	N	Mean	Std. Deviation
My team was open to ideas from all	109	5.7982	1.4258
I was comfortable sharing ideas with team	109	5.8624	1.3015
Team members were kept informed	109	5.1009	1.6327
Collaborative Technology made it possible for my ideas to be understood	109	4.7615	1.4136
Collaborative Technology is an efficient way to work in distributed environment	109	4.7982	1.6146

PERCEPTION OF INFORMATION QUALITY

The participants rated the quality of information unevenly. Table 10 shows the questions and the mean scores. The lowest mean scores were recorded for information accessibility (M=2.74), relevance (M=2.86), accuracy (M=3.27) and understandability (M=3.86). These questions were rated on a 7-point scale (1=low, 7=high). These low scores, when compared to higher scores for sufficiency (M=7.59) and completeness (8.4) suggest that the amount of information was less of a problem than was the ability of users to make sense of it.

Table 10 Perceptions of information quality

Variable	Survey Question	df	Mean Square / Standard Deviation	F	Sig.
INFO1	Information was accurate	6	3.278/1.290	2.542	.025
		102			
INFO2	Information was appropriate	6	4.551/1.504	3.026	.009
		102			
INFO3	Information was accessible	6	2.743/1.969	1.393	.225
		102			
INFO4	Information was relevant	6	2.865/1.586	1.806	.105
		102			
INFO5	Information was timely	6	4.500/1.800	2.501	.027
		102			
INFO6	Information was complete	6	8.401/1.777	4.727	.000
		102			
INFO7	Information was sufficient	6	7.589/1.866	4.068	.001
		101			
INFO8	Information was concise	6	5.544/1.903	2.914	.012
		101			
INFO9	Information was interpretable	6	4.766/1.793	2.659	.020
		101			
INFO10	Information was understandable	6	3.863/1.739	2.222	.047
		101			

The ANOVA showed that not all groups perceived information quality equally. The Command Group consistently rated this element higher than EBAO staff groups. Significant differences were uncovered for questions 1, 2, 5, 6, 7, 8, 9, and 10. Table 12 displays these differences for those questions. As noted above, the staff was in general agreement that information was not accessible (question 3, M=2.74) or relevant (question 4, M=2.86). This finding that the Command Group rated the quality of information they received as higher than other groups is not surprising given the fact that information presented to them was first filtered by the staff. We can be interpret this to mean that the staff did a good job of filtering out information that was incomplete or insufficient prior to presenting updates to the commander, while at the same time noting that the Knowledge Management elements in place did not function as intended; that is to aid in the fusion of information across the staff.

In Table 11 we see that the MNIG and EBE groups rated the quality of information for most categories lower than other staff groups. This is consistent with the EBE problems from gaining information of sufficient detail from component commanders.[7] The problem of information quality with respect to the MNIG may be related to connectivity problems experienced by that group. At times during the experiment, the MNIG team had trouble using the collaboration software and could not access information that was shared with the headquarters staff. In Table

[7] Throughout the experiment, the EBE group noted that they did not have the correct level of detail from the component commanders (played by confederates) to do a good job of operational level execution. This was an experiment design problem that was difficult to adjust once the experiment began.

12, the group identified in column two ('reference group') had a significantly different mean score than the groups identified in column 4 ('comparison groups'). Groups not included in column four had scores that were not statistically different than the column 1 group.

Table 11 CTF Differences between EBAO groups in perception of information quality

Question	Reference Group	Mean	Contrast Group	Mean	Significance
1	Command Group	6.0	EBE	4.11	.047
Accurate			MNIG	3.91	.034
2	Command Group	6.0	EBE	2.37	.012
Appropriate			MNIG	2.18	.045
			SOSA	2.27	.032
5	Command Group	6.25	EBE	3.21	.030
Timely			MNIG	2.73	.004
6	Command Group	6.25	EBA	3.64	.018
Complete			EBE	3.21	.001
			KM	3.95	.034
			MNIG	2.73	.0005
			SOSA	3.82	.036
7	Command Group	6.25	EBE	3.42	.005
Sufficient			KM	3.95	.042
			MNIG	3.0	.002
			SOSA	3.6	.023
8	Command Group	6.25	EBA	3.73	.036
Concise			EBE	3.53	.009
			KM	3.95	.045
			MNIG	3.18	.004
			SOSA	3.60	.026
9	Command Group	6.50	EBE	3.95	.013
Interpretable			MNIG	3.91	.021
			SOSA	3.90	.023
10	Command Group	6.5	EBE	4.16	.027
Understandable			MNIG	4.18	.050

TEAM PROCESS

The Command Group, MNIG, and EBA groups reported high levels of behavior geared toward assigning roles and responsibilities. The SOSA, EBP, KM, and EBE groups had lower ratings for this measure. The pairwise comparison results for the team process survey suggested that significant differences existed only for question 3 (my team was effective in assigning responsibilities). These results are shown in table 12. Though the results for questions 1, 2, and 4 originally suggested significant differences, the strength of the significance level was insufficient for follow-on comparisons. Analysis showed that though all groups rated their team fairly high on these questions, several groups did have lower scores. The EBE and EBP groups were the largest groups and possibly struggled with team process due to that fact. The SOSA

was a smaller group, but because of the nature of their work product, tended to work in isolation rather than together, possibly arguing for a reduced need for group process. [8] The KM group also tended to work in semi-isolation, fixing technical and computer problems as staff asked for help.

Table 12 Team Process results

Variable	df	F	Sig.
My team was effective in sharing information	6	2.477	.028
	98		
My team was effective in assigning roles	6	2.688	.019
	98		
My team was effective in assigning responsibilities	6	3.208	.006
	98		
	104		
My team was effective in communicating ideas	6	2.347	.037
	98		

The significant difference between EBAO groups for question 3 occurred between the Command Group and EBE. The mean for the former was 7.0 compared to the EBE mean of 4.32. This difference between groups was significant at p=.013. This suggests that the EBE group, to a lesser extent than other EBAO groups in MNE 4, did not effectively assign responsibilities to their team members. This could be a result of the leader of the EBE group arriving at the experiment several days late due to uncontrollable problems. However, this points out an important reality for large staffs; that ad hoc leaders must be able and willing to take over for absent leaders.

ROLES AND RESPONSIBILITIES

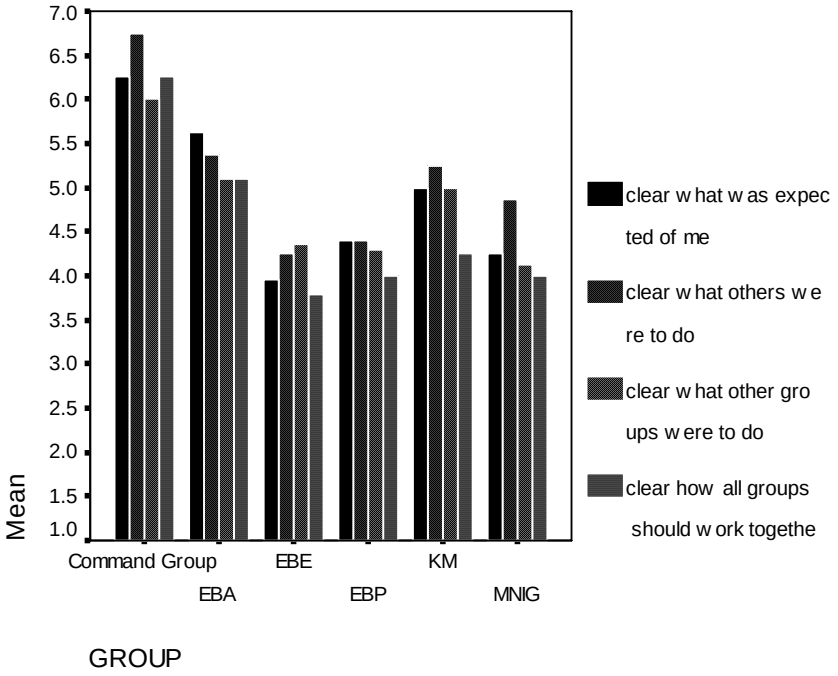
The groups differed significantly on questions 2 (It was clear what others were to do) and 4 (It was clear what other groups were to do) (see Table 7). Figure 2 illustrates that the EBE and EBP groups had lower overall scores, while EBE, EBP, KM, and MNIG demonstrated difficulty knowing what other groups were to accomplish. This is likely related to an understanding of the Concept of Operations for MNE 4 and pre-experiment training, but would certainly be an area of concern for an actual staff.

[8] The SOSA group was comprised of PMESII analysts. When a staff member needed political advice, they would contact the political analyst, who would respond.

Table 13 Between groups differences

Question		Sum of Squares	df	Mean Square	F	Sig.
1. I was clear of what was expected of me	Between Groups	33.572	5	6.714	1.972	.093
	Within Groups	251.978	74	3.405		
	Total	285.550	79			
2. It was clear what others were to do	Between Groups	29.548	5	5.910	2.445	.042
	Within Groups	178.840	74	2.417		
	Total	208.387	79			
3. It was clear what other groups were to do	Between Groups	16.686	5	3.337	1.494	.202
	Within Groups	165.264	74	2.233		
	Total	181.950	79			
4. It was clear how all groups should work together	Between Groups	30.383	5	6.077	2.372	.047
	Within Groups	189.567	74	2.562		
	Total	219.950	79			

Figure 2 CTF results roles and responsibilities



Discussion

In this examination of team collaboration, we sought to identify factors that could explain the performance of a multinational coalition staff. Our analysis provides several opportunities for staff improvements that are uncomplicated and easy to address.

The staff group examined in this study had a mix of large and small groups, both in terms of size, nationality represented, and background. The staff was primarily a military group (with 86% reporting prior military experience) and an experienced one at that (84% of that category had 16 or more years of experience). However, this experience did not extend to the multinational domain. 86% of the experienced participants had less than three years of experience in this area. This lack of experience may partially explain the differences in workload identified during the experiment. In particular, this lack of familiarity in the multinational domain can explain the high levels of frustration experienced. During the experiment, many observations documented the misunderstanding that often occurred between the military and diplomatic staffs, primarily in language and staff processes.

Workload measures are a valuable tool for the measurement of organizational health and should be used with other relevant measures. The NASA TLX is a widely used tool to measure six components of workload, it is easily administered survey, takes very little time and effort to complete, and has excellent validity.

In this experiment, the Command Group consistently rated satisfaction with their own performance higher than the subordinate staff groups. This could be due to the fact that the command tasks in the MNE 4 execution of EBAO did not change from traditional military standards. This could not be said of the staff groups that were required to perform very different tasks, using new and unfamiliar technology, and executing military operations with a new language and with nonmilitary participants acting as team members. The question of how the Command Group tasks, activities, and interactions should change to support EBAO should be explored in detail in future experimentation. Specifically, the virtual presence of the command staff in the operational process should be explored.

It is important to note that frustration between groups was significant in the first and second week of the experiment. In week one, EBP had a higher frustration level than EBA or MNIG. The MNIG group also reported lower frustration than EBE, KM, Red/Green Team, and SOSA. These ratings are supported by observations that showed the EBE and EBP groups working diligently to understand their respective roles in the experiment. These groups also struggled frequently with new software programs designed to improve the planning and execution process. These tools frequently require significant time for users to develop expertise and, as a result, most users were unskilled. In many cases, one or two team members were assigned the job of mastering the tool. Though training was provided before the event, this was not sufficient to master all tools.

The lower frustration levels for MNIG in week one are likely related to the slow formation of this group and their belated understanding of the larger experimentation issues. The MNIG frustration levels rose in the second and third weeks. Reviews of observations suggest that the non-military participants found it difficult to integrate into the military process which

they perceived as rigid, fast-paced, and difficult to understand. In many instances, the MNIG members, operating in distributed teams in the ratio of 'one to many', reported that their opinions were drowned out or marginalized. On other occasions, MNIG representatives were involved in military activities that were inappropriate. [9] This mistake became a valuable learning opportunity in the experimental setting, but speaks again to the basic need to share knowledge of responsibilities among the entire staff.

The staff teams in this EBAO headquarters were able to develop trust in their team and with the technology in use. The highest response scores from the EBAO participants in this survey were for the questions "My team was open to ideas from all" and "I was comfortable sharing ideas with my team." These mean results were 5.79 and 5.86, respectively. These questions were measured on a 7 point scale, suggesting that these scores are high. Lower scores were reported for the questions that asked if the collaborative technology was sufficient to have ideas understood and was an efficient way to work in a distributed environment. These mean scores were 4.76 and 4.79, respectively. These ratings suggest that the collaborative technology, though occasional problems were experienced with the network connection, was accepted by participants and should be pursued as a future work technique.

Information quality remains a difficult problem for an operational headquarters conducting EBAO. This is in part an information technology issue and also a human perception issue. Advanced technologies are needed to distribute information appropriately. This is the problem of getting the *right* information to the *right* people at the *right* time in the *right* format. The fact that the Command Group consistently rated their perception of information quality they received higher than EBAO staff groups likely represents the information fusion capability performed by the staff prior to sharing information with their leaders. And finally, respondents reported that information was complete and sufficient, but less timely, understandable, and accurate. The routing of information through a labyrinth of software programs and teams was the likely cause of this delay and confusion. One future response to information overload is the better articulation of roles and responsibilities.

The EBAO groups identified by the survey as having the most clearly defined roles and responsibilities included the Command Group, MNIG, and EBA. These were small groups that had clearly defined and a narrow range of tasks. The larger groups that had more tasks struggled with assignments. Over time, these groups successfully formed informal sub-groups that seemed to function more efficiently. The MNIG, though a new concept for the military, represented a homogenous group of interagency experts who understood the problems of humanitarian operations extremely well. This group, while not familiar with the military staff and structure, did exhibit a solid understanding of their purpose in MNE 4. The EBA group also represented a small group that was led by an extremely competent individual with vast experience in MNE 4 and previous experiments in this series. This leader took time in pre-event exercises to detail the roles and responsibilities of the EBA. This work was evident in the survey responses reported here.

[9] On one occasion, an MNIG member was involved in a time-sensitive targeting decision (whether or not to destroy a target).

Conclusion

The surveys used in this experiment were helpful in understanding the uneven process of developing teamwork and trust in the conduct of this distributed collaborative event. It is noteworthy that the teams, each of which was comprised of members from the participating nations, reported that team members established a trusting relationship. In spite of this, the bulk of the staff did not report effective team processes. As a result, much of the staff did not fully understand the inter-team relationships and how all teams would work together to achieve overall goals. This confusion was evident in the execution of the process, and most notably, in the integration of the MNIG into military tasks. The lesson learned from this experiment is that non-military actors cannot simply be added onto a military operation like supplemental staff. Their contributions and their viewpoints are unique, and they must be integrated into the staff with careful consideration for their contributions. All staff must understand the roles and responsibilities of other groups.

The lack of shared knowledge of overall roles and responsibilities can partially explain the unequal distribution of effort, performance, and frustration and the unmanageable sharing of information. Larger teams reported increased workload, lower performance, and higher frustration. Respondents' ratings of information quality suggest that information was complete and sufficient, but less timely, understandable, and accurate. This suggests that information was pushed to many, regardless of function, rather than pushed to selected teams or individuals.

The frustration element appeared to be generated by two primary factors that were not shared equally among the teams. The interagency group (MNIG) expressed significant frustration at their lack of understanding and familiarity with the military staff process. They considered it rigid and complex. Military groups, on the other hand, expressed frustration at the software tools that had been designed to support planning, execution, and assessment activities. These tools were designed to help the staff deal with large amounts of information and to provide a visual model of the progressing operation. Software tools to support knowledge management in a coalition staff continue to mature. This evolution must be supported by ongoing human-system integration research to ensure that users can use these tools and reduce their cognitive workload. There is clearly much work left to be done in this area.

These findings suggest that smaller teams were able to establish the prerequisite characteristics needed for performing in a capable manner. System designers should consider maximum group sizes in staffing military headquarters. This is especially important when the primary communication device is a collaborative software system and the team members will never meet in a face to face medium. In this environment, training is an essential and critical skill that is often unappreciated by designers and users. Many of the skills used in MNE 4 were quickly lost if not used routinely. The training staff should continue to work hand in hand with concept developers to sequentially build a training program. This would allow for a 'just in time' approach to training that could be instituted at each stage of the experimentation planning process, or a staff building process. Training for EBAO in a collaborative environment is substantial. Not only are there explicit learning objectives, but implicit objectives as well. Explicitly, all staff members must learn their own and overall roles and responsibilities. The implicit objectives speak to the technological and social developments that argue for EBAO.

References

Edmunds, A. & Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management* 20: 17-28.

Fine, G. A. & Holyfield, L. (1996). Secrecy, Trust, and dangerous leisure: generating group cohesion in voluntary organizations. Social *Psychology Quarterly*, 59 (1): 22-38.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of four subjective workload rating scales. Human Factors: 34 (4), 429-439.

Mayer, R. C., Davis, J. H., & Schoorman, E. D. (1985). An integrative model of organizational trust. *Academy of Management Journal* 20(3): 709-734.

Smith, E. (2002). Effects Based Operations: Applying Network Centric Warfare in Peace, Crisis, and War. Washington: CCRP.

Stevens, J. (1996). Applied multivariate statistics for the social sciences. Mahwah, N.J.: Lawrence Erlbaum Associates.

U.S. Joint Forces Command. (2006). The Multinational Experiment 4 Effects Based Operations Concept of Operations. Suffolk: author.

Warne, L., Ali, I., Bopping, D., Hart, D. & Pascoe, C. (2004). The Network Centric Warrior: The human dimension of network centric warfare. Edinburgh South Australia: DSTO Information Sciences Laboratory.