Objectively Measuring the Promulgation of Commander's Intent in a Coalition Effects Based Planning Experiment (MNE3).

Andrew Leggatt, BAE Systems,
andrew.leggatt@baesystems.com

Acknowledgements

The author of this paper would like to acknowledge the very large contribution of others to undertaking this work, without whose input the paper and the task would not have been completed. These include: Nigel Paling, Iain Saker, John Coote, John Admire, Chris Elliott, Mike Kalms, Martin Pepper, Simon Atkinson, Jeremy Wilson, Andrew Preece and Steve Potter.

"Never tell people how to do things. Tell them what to do and they will surprise you with their ingenuity."
- General George S. Patton, Jr

Abstract

From the beginning of written warfare history, such as Alexander defeating Darius (the Persian King) at Issus in 333BC, war fighting has been characterised as the Commander imposing his will upon both his own and enemy troops to achieve his desired outcome (Keegan, 1999). These basics of war fighting have not changed over the millennia and it is the exertion of the commander's will that is now written into most nations' doctrine as "commander's intent". Without the Commander successfully communicating his intent to his subordinates there is little chance that the overall mission will be a success. It is thought that with the rise in Effects Based Operations (EBO) that commander's intent will become more important as NEC will provide subordinate commanders greater room for manoeuvre whilst there are less onerous control restrictions.

This paper provides a summary of work undertaken by NITEworks for DCBM/J6 to quantitatively evaluate how effectively the Commander's intent has been promulgated through a simulated, distributed-Coalition Task Force HQ (CTFHQ) during Multi-National Experiment3 (MNE3). The Signal Detection Theory (SDT) paradigm was used to evaluate the understanding of commander's intent. It would appear that commander's intent was not dispersed beyond the (centrally located) Command Team (comprising of Commander, Deputy Commander, Chief of Staff, and chiefs of each group), the CIACG (Coalition Interagency Co-ordination Group) and the Plans team. This may be of concern to policy makers since NEC is credited (House of Commons Defence Select Committee, 2003) with the potential to increase the understanding of commander's intent and there is a requirement for an improved understanding of the Commander's intent for Effects Based Operations. These findings should be tempered with the caveat that these data represent only the very first steps in trying to quantitatively measure the

Commander's intent and they were collected in an experimental environment using a coalition type command structure.

A number of other conclusions are drawn which include that the method appeared to successfully examine commander's intent, non military personnel appeared to understand it and the understanding of commander's intent was influenced mostly by organisational proximity to the Commander.

Background

This paper describes work undertaken by NITEworks (Network Integration Test Experimentation works which is a MoD/Industry partnership which assesses the benefits of NEC through experimentation) for DCBM/J6 to quantitatively evaluate how effectively the Commander's intent has been promulgated through a simulated, distributed Coalition Task Force HQ (CTFHQ) during Multi-National Experiment3 (MNE3). The paper describes measuring commander's intent using the Signal Detection Theory (SDT) paradigm to provide a quantitative indication of whether HQ staff comprehend the Commander's intent.

Introduction

From the beginning of the written history of warfare, Sun Tzu circa 500BC or Alexander defeating Darius in 333BC, warfighting has been characterised as the Commander imposing his will upon both his own and enemy troops to achieve his desired outcome (Keegan, 1999). These fundamentals of warfighting have not changed over the millennia and it is the exertion of the commander's will that is now written into most nations' doctrine as "commander's intent". Without the Commander successfully communicating his intent to his subordinates there is little chance that the mission will be a success. It is thought (House of Commons Defence Select Committee, 2003) that with the advent of Effects Based Operations (EBO) that commander's intent will become more important as NEC (Network Enabled Capability) will provide subordinate commanders greater room for manoeuvre whilst there are less onerous control restrictions.

Commander's intent is defined by JWP 0-01.1 as, "A concise expression of the purpose of the campaign or operation, the desired results and how operations will progress towards achieving the desired end-state. At the tactical level, the commander's intent should be focused on the effect that he wishes to achieve on the enemy". This definition suggests that the epitome of command intent is the end state required and not what the troops have do to achieve this and, as such, may be seen as the cornerstone of the operational art.

Commander's intent is a key concept to UK doctrine and it marries two other key aspects of UK military doctrine, the manoeuverist approach (which seeks to gain advantage over the enemy without necessarily engaging in attrition) and mission command (mission led orders which direct subordinates as to what is to be achieved and why, but not how. Both aspects were formally introduced to UK doctrine in the 1980s). These concepts empower the UK military to punch much heavier than their numerical strength would suggest.

The essence of these concepts are that of freeing the Commander on the ground, within bounds defined by his superior command, to have the freedom to manoeuvre and take action to complete his mission in the way he thinks most suitable. In summary, the subordinate Commands and all staff and enlisted men and women need to understand the purpose of the activity so they can apply themselves with a *unity of effort* whilst using their initiative to undertake this in the most efficient way.

The early advocators of mission command or "Auftragstaktik" was the post WWI German army who developed their doctrine in a landmark document, Truppenfhrung (1933 and 1934). Interestingly, the term "Auftrangstaktik" does not appear in the manual but makes statements such as, "The general intent must be stated for the execution of impending operations, but the method of execution is left to the subordinate commanders. Otherwise it becomes a directive." (Truppenfhrung, para 76, 1933 – translation Condell & Zabecki (2001)). This is clearly "mission command" and requires an unambiguous enunciation of the commander's intent.

Pigeau and McCann (2000) have suggested that commander's intent should be redefined as "common intent". They argue that there is a need for the subordinate commanders to understand not just their commander's explicit intentions, as written or vocalised during the Orders process, but also to understand the implicit meaning of the intent. For example, the Commander may want the mission to capture "Dog Wood" but would implicitly expect his troops not to harm the civilians in the farmhouse adjacent to the wood. This is a fair suggestion but many commanders would expect their implicit requirements to be met and undertake many practical activities to ensure they have a shared understanding of the "problem space" to verbalise the most salient aspects of the implicit tasks. For example, Nelson requested that the Captains of the English flotilla, preceding the Battle of Trafalgar, dine with him each evening to share his understanding of the battle space (Alberts and Hayes, 2003) and make it common. Therefore in this paper the term "commander's intent" will be used as a readily understood concept which is written into UK doctrine.

The importance of commander's intent to EBO was highlighted by the Joint Doctrine and Concept Centre (JDCC) and is articulated in the New Chapter of the Strategic Defence Review as, "We are looking at how mission command in the information age can work. We are agreed that mission command encapsulates one of the best aspects of the British approach to the use of military force, the ability for a commander to articulate his intent and for the people beneath him to decide on the best way of carrying that out. The information age should allow a much greater dissemination, a much clearer exposition, of the commander's intent." (Air Vice Marshall McNicoll, pp 34, House of Commons Defence Committee, 2003).

The experimental environment where the measurement of commander's intent was exercised was Multi-National Expeirment3 (MNE3). This was an Operational Level of Command planning experiment which used a new Effects Based Planning (EBP) process developed by the muliti-national group

with participants from all 6 Multi-Interoperability Council (MIC) nations and NATO (North Atlantic Treaty Organisation). The data reported in this paper were collected in a distributed CTFHQ (Coalition Task Force HQ) with participants sited in each of the separate nations (Australia, Canada, France, Germany, UK and US). However, some elements of the Command and Staff were co-located: the Commander, Deputy Commander, Chief of Staff and each leader of the primary teams within the new organisational structure being practised (Plans, Operations (Ops), Information Superiority (IS), and Knowledge Management (KM). Other teams included CIACG (geographically distributed), Logistics (geographically distributed) and Medical (centralised within the USA)). Each team, however, comprised staff from all participating nations i.e. from Australia, Canada, France, Germany, UK and USA such that each nation's site had a mix of participants in all of the teams.
The method used to evaluate commander's intent was a development of the QUASA™ method (McGuinness, 2004). This method uses probe statements about the situation and requires participants to judge whether they are true or false. Their responses can be interpreted using the Signal Detection Theory (SDT) paradigm which allows hits, misses, false alarms, and correction rejections to be calculated (see Figure 1: Interpretation of Responses within the Signal Dection Theory Paradigm below). This provides a performance measure but also gives an insight into the bias of the participants. For example, it is possible to ascertain whether groups or individuals are more likely to agree with false information or reject true information.

	Assessed as "Correct"	Assessed as "Incorrect"
Actually Correct	**HIT**	**MISS** (error)
Actually Incorrect	**FALSE ALARM** (error)	**CORRECT REJECTION**

Figure 1: Interpretation of Responses within the Signal Dection Theory Paradigm

Method

Design

A quasi experimental design was adopted for MNE3. No independent variables were manipulated but a series of naturally occurring independent variables existed (type of tasks ongoing in the HQ (also referred to as "Date"), geographic distance from the Commander ("Site"), and functional group (IS,

KM, Ops, Plans, etc). This provided a 4 (dates or HQ tasks: Effects Assessment, Action Assessment, COA war gaming and Synchronisation) x 8 (groups: Command group, Plans, Ops, IS, KM, CIACG, Log & Medical) design. The date/HQ tasks variable was a within "subjects" variable.

The dependent variables reported in this paper are task performance to the probes (hits, misses, false alarms and correct rejections) which are converted into d' and β and confidence ratings.

A number of hypotheses were developed. These were as follows:

1. H_1: The method will provide a clear insight into the promulgation of commander's intent throughout the CTFHQ simulated in MNE3.
2. H_1: Those staff geographically closest to the Commander will have higher levels of commander's intent than those further away (i.e. US site greater commander's intent than UK, Australia, Canada, etc.).
3. H_1: Those staff organisationally closest (i.e. the Command team and to a lesser extent the plans team) to the Commander will have higher levels of commander's intent.
4. H_1: The probe performance on different days/events will provide an indication of whether there were problems with the Staffs' understanding of commander's intent.

Assumptions

Participants' confidence ratings will provide an indication of task difficulty. If there is no significant difference in probe confidence between days, it can be assumed that the questions are of equal difficulty.

Participants

There were 115 participants in the MNE3 experiment and 105 completed all 4 sets of commander intent probes (5 of the participants who did not complete all the questions were from the UK, where team members interchanged because they could not be available the entire period of the experiment). The participants in the experiment were mostly serving military, or recently retired[1] military personnel, from the six MIC nations, apart from those 8 civilians in the CIACG who were undertaking special support roles (e.g. Foreign and Commonwealth Office, DfID[2] etc.). In general the participants had a large amount of military experience with the median experience being 16-20 years (see Figure 2: Experience of MNE3 participants). The participants were mostly from operational planning backgrounds and the UK participants comprised the JFHQ (Joint Force HQ) operational planning team from PJHQ, Northwood.

[1] US contributors to the experiment were mostly retired officers contracted to USJFCOM
[2] DfID (Department for International Development)

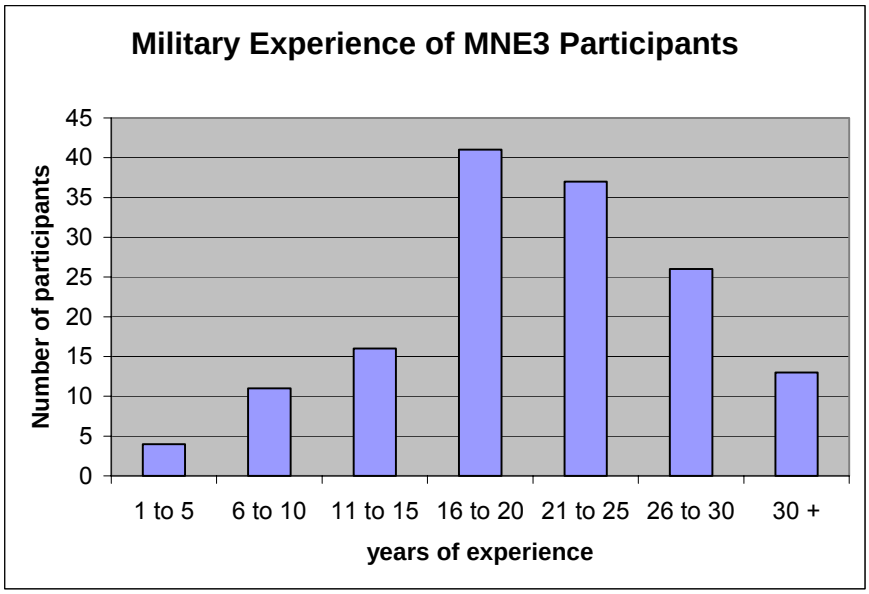

Figure 2: Experience of MNE3 participants

The participants were grouped into Plans, Ops, IS and KM with a number of supporting additional roles. These included the CIACG, Logistics & Medical. For the interests of the experiment the command team was considered to be the Commander (retired USMC MG, deputy Commander (retired German MG), COS (Chief of Staff), and the leader of each of the four primary teams. This grouping was partly due to the understanding that UK had of the organisation but partly it provided a more equal number of participants in each of the groups for sensible comparisons. The Commander did not answer the probes since he set their content.

The number of participants who completed all 4 sets of experimental probes are listed in the groups is shown in Table 1: Number of participants by role.

GROUP	NO. OF PARTICIPANTS
Plans	20
IS (Information Superiority)	45
Command Team	6
CIACG	8
OPS	11
Logistics	3
KM (Knowledge Management)	5
Medical	7
All Groups	**105**

Table 1: Number of participants by role.

The number of participants by role was dominated by IS and Plans because the HQ was constructed to address the planning part of the HQ function, whilst it was not required to conduct the operations phase.

The participants were distributed over the six MIC nation sites with the command team resident in the US. The distribution of participants is shown in Table 2: Number of Participants based in each country.

NATION	NO. OF PARTICIPANTS
Canada	11
US	42
UK	15
France	15
Australia	6
Germany	16

Table 2: Number of Participants based in each country

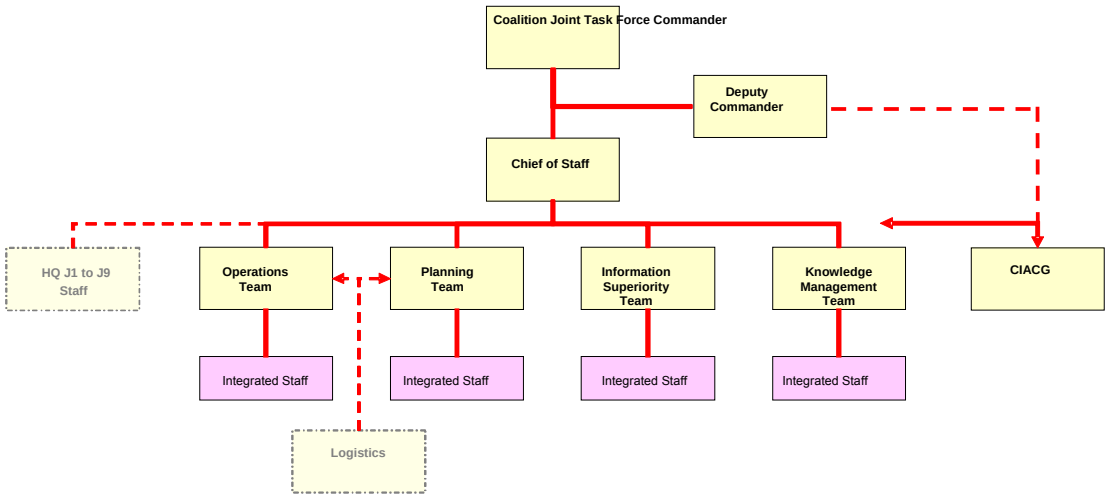

Figure 3: Organisational diagram of CTFHQ

Materials

The materials used to undertake this experiment were dominated by IT (Information Technology). Since the experiment was conducted remotely at different sites the IT had to support both group working (teams of people working remotely together) and task working (particular special activities).

The key group-working tool was IWS (Information WorkSpace) which "connected" participants and allowed them to text chat, voice chat, file share, attend lectures, have discussion threads, application share and use a white board. This software relied on a federated server network with standard range PCs networked on a classified WAN (the Combined Federated Battle Laboratory Network or CFBL Net).

The task working tools were sited on each of the networked PCs and included the usual MS 2000 operating system with MS Office and in addition a bespoke software application developed by QinetiQ, under contract with the MoD, to support EBP at this level. This tool, commonly referred to as the "Q" tool, provided an integrated environment for the participants to undertake the EBP task for each of the different stages, recording the decisions and work done and making it available for the next task step.

To collect the commander's intent data, a software package called JDCAT (JBC Data Collection Analysis Tool) was used to push probes to each of the experimental participants. JDCAT collected the responses and provided a "delinquency" report of who had not completed probes allowing experimenters to remind participants to complete their assignments.

Procedure

The procedure for measuring the promulgation of commander's intent can be divided into two parts, the construction of the material and the administering of the material to the participants.

The probe materials were developed in close association with the Commander, the (UK) Plans Team Leader and advice from a retired UK General who was one of the UK's Senior Concept Developers. The core or base probe statements were constructed from the Commander's published "Commander's Initial Guidance". These probe statements were constructed with reference to a number of "wicked questions" or extremely testing problems for coalition commanders in the field suggested by the UK Senior Concept Developer. Whilst constructing these probe statements a rubric of requirements was identified.

The probe statements shall:

1. Reflect the commander's intent,
2. Use simple language (probes should not be a language test for non-native English speakers)
3. Not be a verbatim copy of the published guidance
4. Be a mixture of implicit and explicit issues.
5. Be operationally relevant.
6. Be equal in number true and false.

An example probe statement is shown below:

The Commander will escalate measures against uncooperatives (true).
CTF considers that food aid will be controlled to exert influence (false).

Figure 4: An example of two probe statements

The base probe statements were then saved in a password protected file and sent to the UK scientific liaison in the US for briefing to the Commander. The Commander scrutinised the probes and modified them where necessary to reflect his view of his intent. A number of changes to the probes were made and some probes were deleted because they were irrelevant or did not reflect his intent.

Then the Commander selected 6 probe statements (always comprising some "true" and some "false" probes, but not necessarily in equal number) from a choice of approximately 15 probes which he thought were the most accurate reflection of his intent and these were then given to the analysis staff.

The analysis staff put the intent probes into the JDCAT tool and these probes were then pushed to all experimental participants in all MIC nations. The probes appeared on the participants' screens when they choose to complete their questionnaires, usually at the end of the experimental day. The participants were instructed not to consult their fellow participants and had to judge whether the probes were true or false and provide a confidence rating on a 1 to 7 scale from extremely unconfident to extremely confident.

Analysis

The analysis of the probe data was based upon the SDT paradigm. The data collected was categorised as follows:

"Hit" (a true statement which the participant agreed was true);
"Miss" (a true statement which the participant thought was false);
"False alarm" (a false statement which the participant thought was true); and
"Correct rejection" (a false statement which the participant correctly rejected).

Once the data were categorised into these sub groups the commonly used statistics for SDT were be applied - sensitivity (d') and bias (β). SDT states that these two measures separate the d' of the individual from their β. The measures suggest that the greater the sensitivity, the better the individual is at identifying the signal (true probe) from the noise (false probes) This can be represented as distance between the distributions (see Figure 5, below).

β is a reflection of where the individual places his/her criteria (see figure 5, below). The participant may choose to accept information which they are unsure about, and therefore minimise the amount of true information they reject and therefore accept most true information. This behaviour is attributed

to be liberal or risk taking strategy. This type of behaviour may be suitable where there is little penalty for accepting false information. This is reflected as a low β score. Alternatively, participants may choose to adopt a "conservative" approach and reject probes which they are unsure about and therefore probably miss more true information. This criteria or bias is shown by a high β value. Such an approach is often considered suitable in situations where there are severe penalties for accepting false information (e.g. criminal cases where the defendant would be executed).

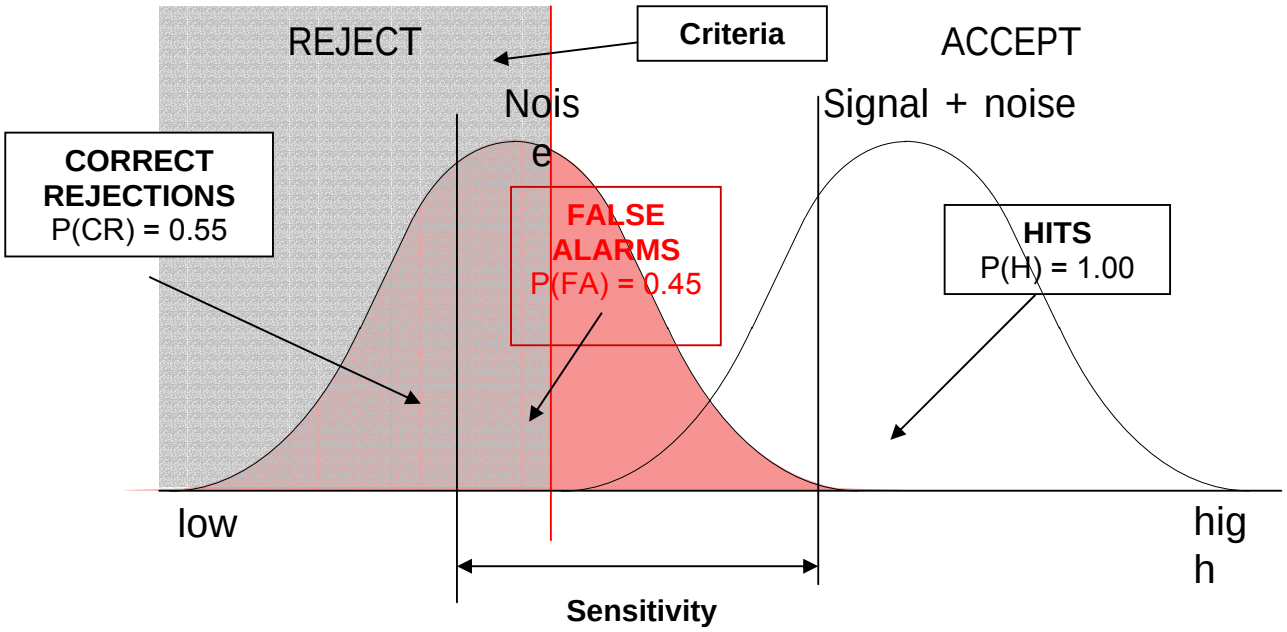

Figure 5: Diagrammatic representation of SDT

The data collected during the experiment were aggregated and mean β and d' created for each player, invoking central limit theorem allowing the data to be treated as interval data and analysed with parametric statistical tests. On those occasions when a participant failed to make a false alarm, (this happened on a small number of occasions) and thus making it impossible to calculate d' the data were modified in line with Wickens (2002). He suggests arbitrarily modifying the FA rate from 0 to 0.01 and likewise reduce the hit rate to 0.99 to allow d' to be calculated.

A mixed 4 x 8 design ANOVA, with repeated measures for date/event, was undertaken to establish whether there were any main effects. When main effects were identified post hoc comparisons were undertaken using Duncan's multiple range test.

The confidence data were analysed separately to the probe data. These data were of an ordinal nature but a mean confidence rating was calculated which allowed central limit theorem to be invoked. These data were analysed used a mixed 4 x 8 design ANOVA, with repeated measures for date/event. Where significant main effects were established Duncan's multiple range test were applied to identify individual differences.

Results

Confidence ratings

Site differences

There was no significant difference between the sites. This suggests that each site found the questions approximately as difficult as each other or that they were equally unable to discern their ability to answer the questions correctly. This supposition is not supported by performance data (see analysis below) which suggests a significant difference in site performance.

A possible absolute observation that can be made of the data is that the ratings are consistently greater than the middle point (4) (mean confidence rating 5.2 with sd 0.94). This would suggest that participants are generally confident about their responses.

Date/event differences

There was no significant difference for events/date in confidence ratings. This may suggest there are no substantial differences between these events on different days.

Functional group differences

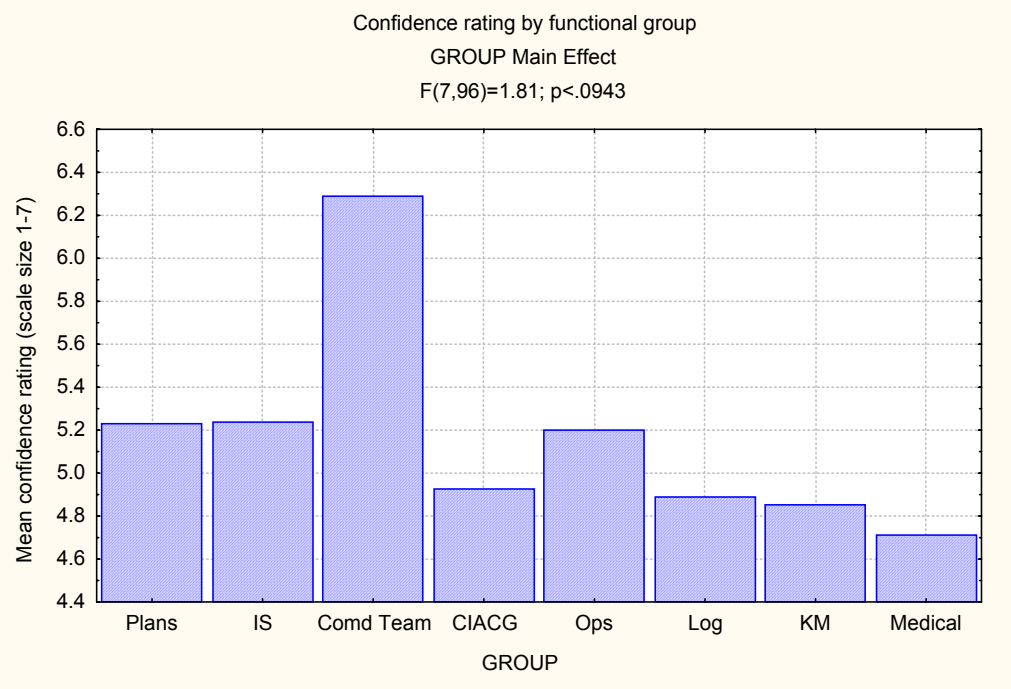

Figure 6: Mean confidence rating by functional group

The ANOVA suggests the differences between functional groups were approaching significance (P<0.10) although not strictly significant by the 5% convention.

Duncan's multiple range test identified that there were significant differences (Duncan's multiple range test is a reasonably liberal test of post hoc differences which is judged suitable for field experimentation data like these) (P<0.05) between the Command Team and all other groups. There were no other significant differences. This suggests that the Command team had significantly greater confidence in their ability to answer the probe questions correctly than the rest of the CTFHQ. This result was expected. The Command team were both close geographically and organisationally to the Commander and therefore should have been in a better position to understand what the Commander wanted to achieve and to be confident about this.

Sensitivity – d'

Site differences

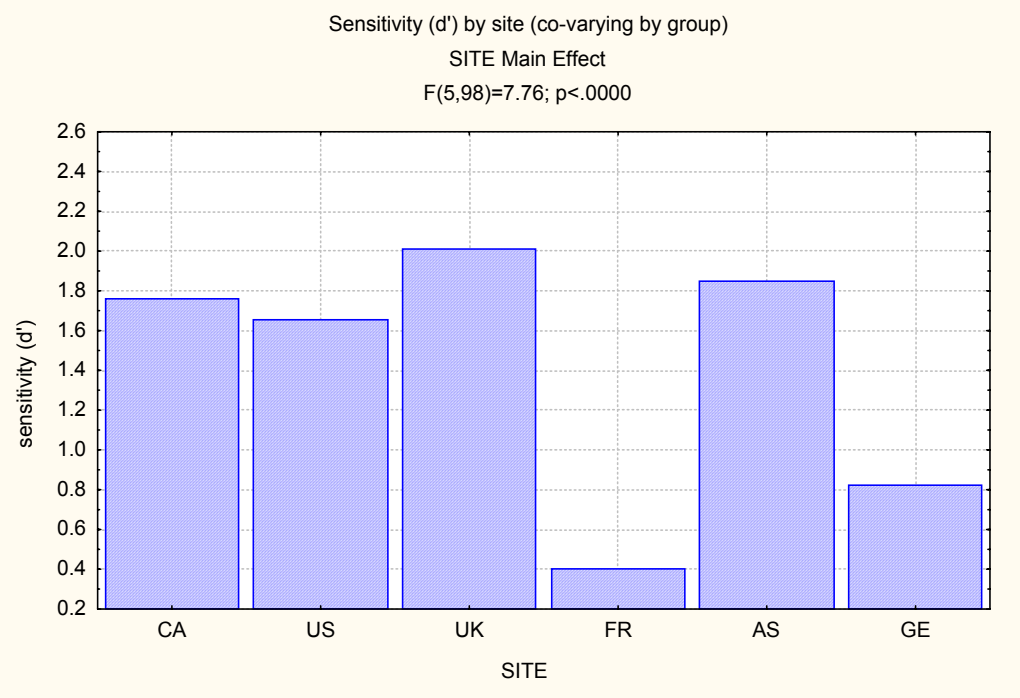

Figure 7: Participant sensitivity (d') to probes by site

The ANOVA identified a significant effect of site $F(5, 98)=7.76$; $P<0.01$. Post hoc comparisons using Duncan's multiple range test identified that there was no difference between the French and German participants' performance but their sensitivity was significantly different to all other nations performance.

These findings suggest that there was a language effect in the probe test. Those CTFHQ participants whose first language was not English and were in remote sites (France and Germany) performed significantly worse than the other nations. The effect is likely to be due to language and not remote location because their performance was significantly different to other remote sites such as Canada, UK, and Australia but whose native language is English.

Functional group differences

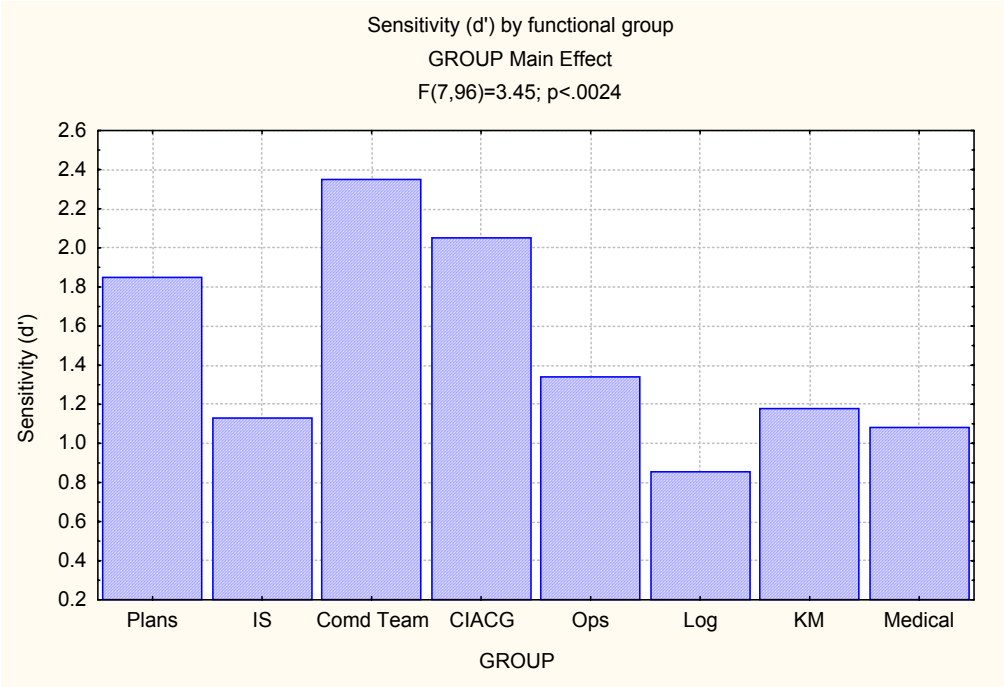

Figure 8: Sensitivity (d') to probe statements by functional groups

The ANOVA identified a significant difference between functional groups $F(7,96)=3.45$, $P<0.0024$ (more variance was accounted for when co-varying by site with F values increasing from 2.6 to 3.45).

Further analysis of the data using Duncan's multiple range test identified a significant difference between the Command team and IS, Log, Ops, KM and Medical. However, there is no significant difference between Plans, CIACG and the Command team. In-fact CIACG is significantly different to Log and approaching significance for IS and Medical.

These data suggest there are 2 groups, those with reasonable sensitivity (Command team, CIACG & Plans) and those with very little sensitivity (IS, Ops, Log, KM, & Medical).

Bias - β

Site Differences

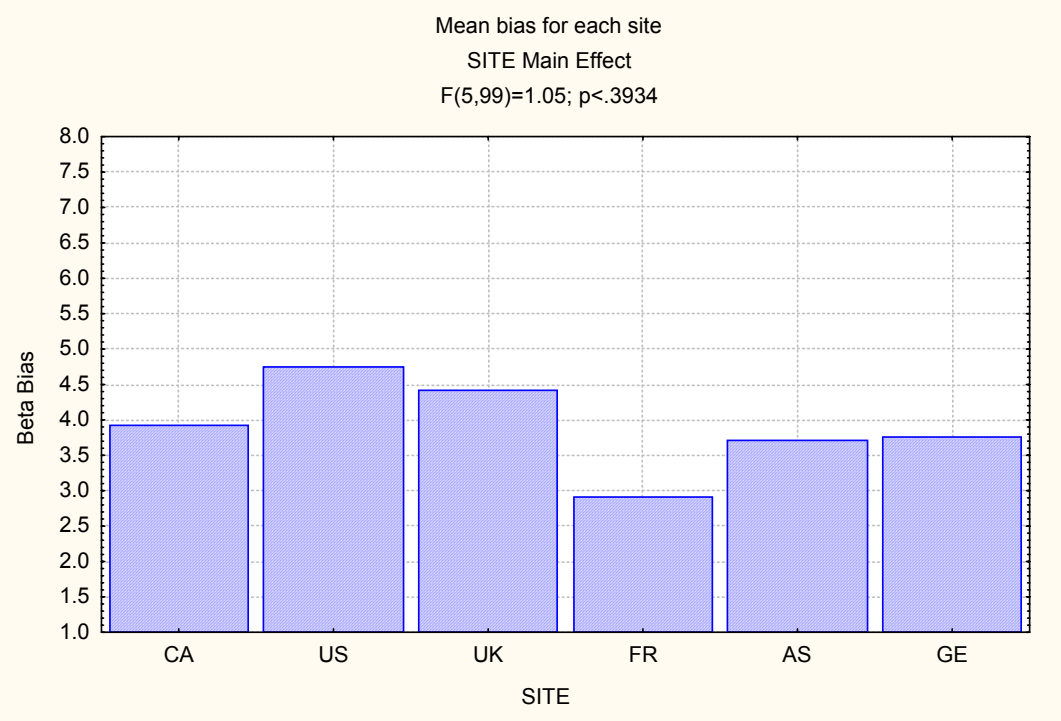

Figure 9: Mean Beta by site

No significant main effects in β or bias were identified between sites. It would appear that none of the participating nations differed significantly in the criteria that they adopted.

Functional Group Differences

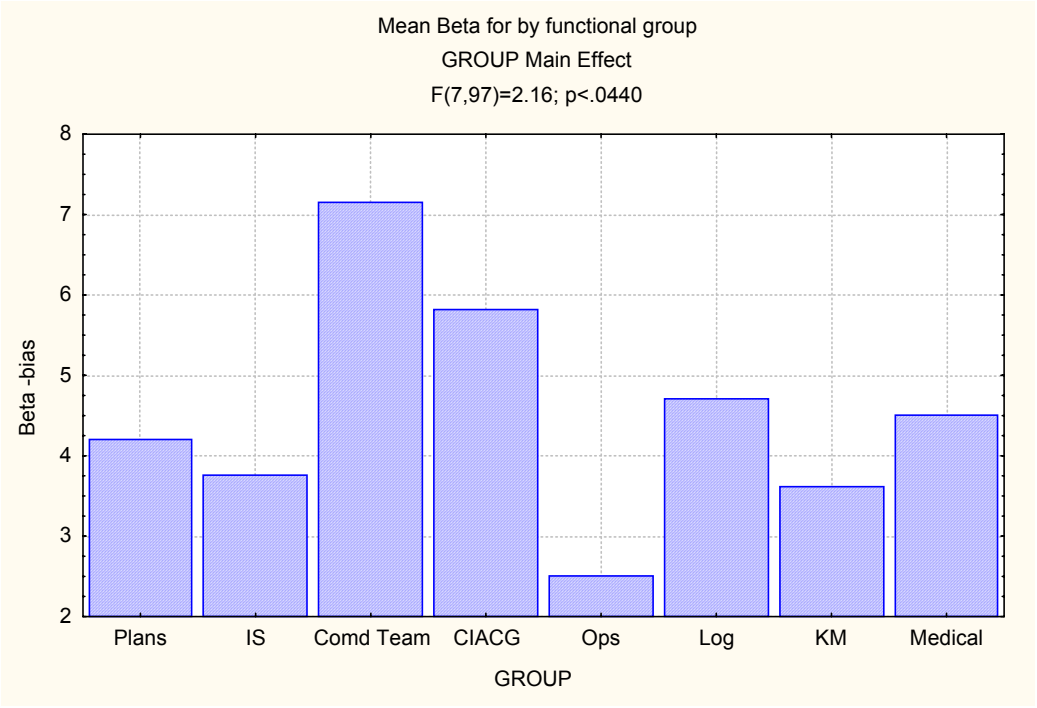

Figure 10: Mean Beta by functional group

ANOVA identified significant main effects between groups F(7, 97)=2.16; P<0.044. Further analysis using Duncan's multiple range tests identified that the Command team had significantly greater β than IS, Ops and KM with approaching significance (P=0.07) for Medical. The CIACG is approaching significance to Ops (P=0.055).

These data are quite interesting as they seem to be related to the d' data which indicate that both the Command Team and CIACG have greater sensitivity, however, bias and sensitivity are independent measures. A change in criteria will not improve sensitivity, it may however optimise performance for the task (i.e. false alarms may be acceptable when trying to identify bombs at airport check-in facilities, whereas false alarms would not be acceptable when falsely convicting innocent men in a court of law.

The Command Team and the CIACG appear to be less likely to accept false information.

Performance results from probability data
Probability of False Alarms

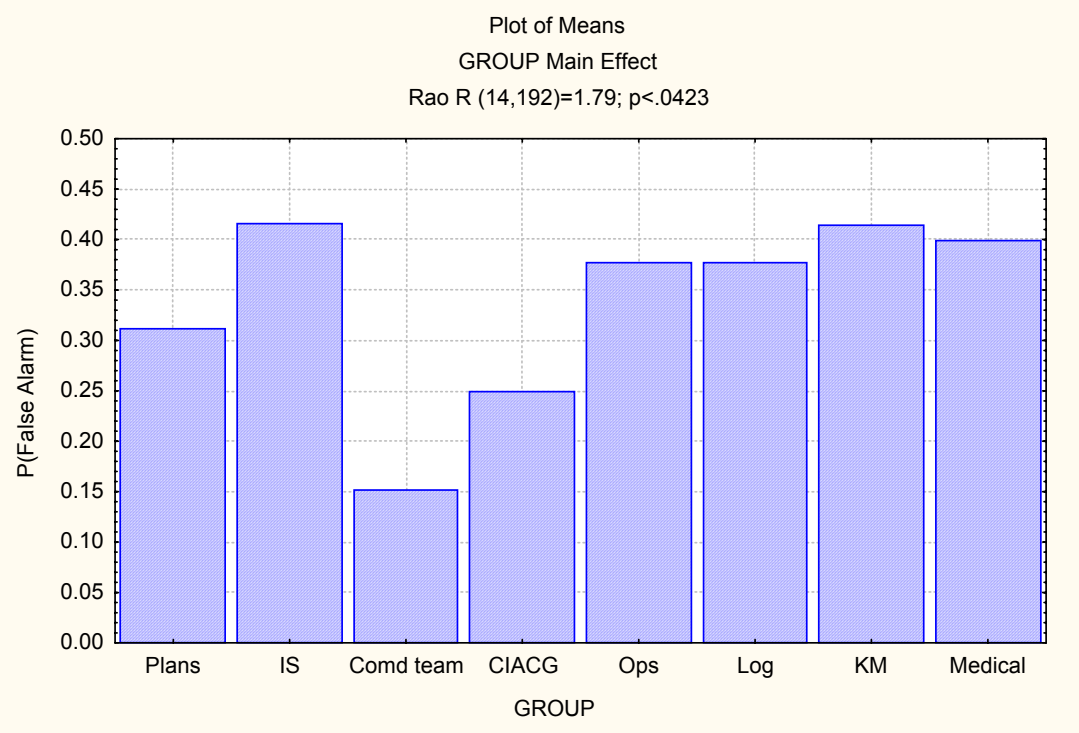

Figure 11: Mean Probability of False Alarms by Functional Group

The Command Team made significantly (P<0.05) fewer false alarms compared to all the CTFHQ apart from the CIACG. The CIACG approaches significant difference to KM. There were no other significant differences.

Probability of Hits

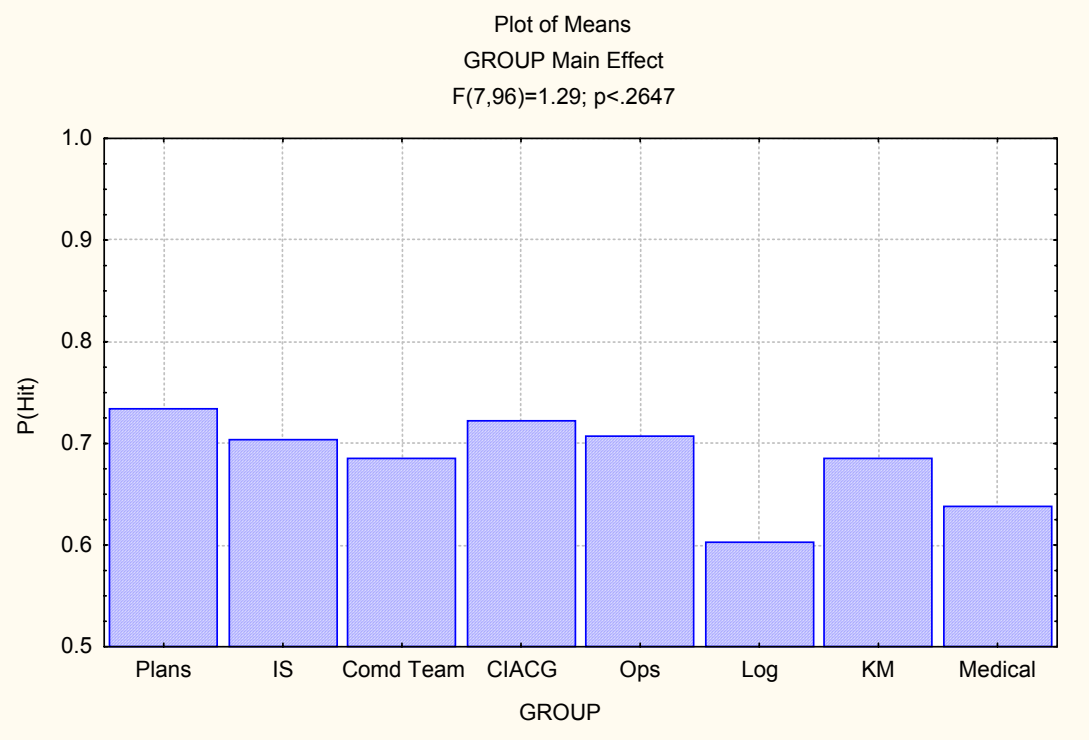

Figure 12: Mean Probability of Hits by Functional Group

There were no significant differences in the P(Hits) between any of the groups.

These data suggest that the ability of the groups to identify true information is similar.

Discussion

The meaning of the results with respect of the hypotheses will be discussed in turn below.

Confidence ratings

The ratings indicate that the level of confidence when answering the probes at different sites and on different dates did not vary. However, there were significant differences between functional groups. The command team had significantly greater confidence than the other groups.

Together these findings suggest that the command team thought they understood the commander's intent to a greater extent than the other elements of the CTFHQ and the objective data supports their confidence. A lack of variation in confidence between sites and dates could be interpreted that these groups experienced no difference in difficulty. That is the German participants found the probes as difficult to answer as the participants at the US site. Likewise, the lack of difference in confidence between dates/events could suggest there was no difference in probe difficulty and that it is possible to compare scores between dates/events and attribute these differences to events that were happening and not to difference in probe difficulty.

However, it would appear that the confidence ratings were not an accurate reflection of task performance for the different sites (e.g. French and German performance was lower than the rest of the CTFHQ but there was no difference in confidence ratings). Therefore these data suggest it is unwise to attribute the differences between dates to the events that occurred because it could be due to differences in probe difficulty. For this reason the difference in date/event will not be discussed further in this paper.

d' – Sensitivity

Site differences

Analysis of d' or sensitivity identified that there was a significant difference in performance between sites. The data suggested there were 2 performance groups, the non-native English speaking sites and the native English speaking sites. This suggests that the probe method identified a language issue in the CTFHQ. The working language of the HQ was English and the Commander's intent was given in English and the probes were administered in English. The poor sensitivity of the Franco-German block could suggest that the understanding of command intent was poorer at these sites **or** that it was an artefact of the test and their understanding was adequate but they could not express it in a test.

This could suggest that the probe test was methodologically weak (i.e. the command intent test is merely a language test) but there are a number of factors which suggest that the data are still useful. These are:

1. The language effect that these data identify is a real one. Those individuals in the CTFHQ who do not speak English as their first language do not exhibit such high levels of sensitivity on the task as the other sites. If this happens in a test it may happen in reality. If this is the case the Coalition Commander need to understand the potential for their intent to be misunderstood and correct for this accordingly. Potentially the Commander needs to take precautions against this misunderstanding.
2. The probes required the participants to have reasonable understanding of English, but the participants also have to have a reasonable understanding of English to understand the Commander's intent and communicate this on in the line of command.
3. During other Multi-National events the author has collected anecdotal evidence that the non-native speakers often prefer the written word to the spoken word because it allows them time to reflect on the language and possibly use automatic translation tools. The probe statements were administered in a written form and there was no time limit on response time. Therefore this should minimise the language effect.
4. Whilst it is regrettable that the non-native English speakers performed less well than the other nations, these data can be used to co-vary the other results and therefore partition the variance in the analysis. This provides a more accurate analysis of the data for the differences between the functional groups.

A possible modification to the experimental methodology would have been to have the probes translated into French and German to determine whether it was a language issue or whether it was an understanding issue.

The other site finding concerns the US site, where the Commander was located but performance was not significantly different to other sites. These data suggest that it is not possible to accept the experimental hypothesis that geographic proximity to the Commander improves the understanding of commander's intent.

There are a number of possible explanations for this finding. The personal style of the Commander meant that he did not exert his physical presence at US site and was often not physically present in the main work room of the CTFHQ (e.g. he was often required to attend closed meetings). Anecdotal evidence based on observations made by the UK scientist in at USJFCOM supports this explanation. In addition, the US site was much larger than the other sites and therefore personal contact between the Commander and his staff was much less likely to occur.

Thus the hypothesis that geographical proximity would increase performance was probably miss-framed. The Commander was probably not appreciably more physically present in the US site than the other sites.

<u>Functional group differences</u>
There were significant differences in sensitivity of the functional groups. The Command Team (excluding the Commander who set the probes), CIACG and

the Plans Team (to a lesser extent) had significantly greater d' than other the groups.

The Command Team were both geographically and organisationally very close to the Commander and would be expected to have the best developed understanding of the Commander's intent. These data suggest that, on the face of it, it is possible to tentatively accept the experimental hypothesis about organisational proximity. Indeed, the Plans Team, who were geographically dispersed, had significantly greater d' than KM, IS, Logistics and Medical who were also geographically dispersed. The high performance of Plans team is inline with this hypothesis.

However, the CIACG team's high level of performance was initially considered surprising. This functional group did not contain any military personnel and were generally not familiar with the military planning process. Possible explanations of this finding could be that they were very able individuals (for example selection criteria for the UK Foreign and Commonwealth Office is extremely demanding) and that they particularly needed to understand the Commander's intent for them to undertake their task of providing non-military input into the military system, whereas other military personnel in other groups had a number of other issues to attend to as well as Commander's intent.

Further examination of the intriguing finding revealed that although the CIACG were not composed of military staff they were organisationally close to the Commander. The CIACG lead (a retired US Ambassador) reported directly to the Commander. This group appeared to substantially influence the Commander's decisions on a number of occasions.

The CIACG having such a good understanding of commander's intent may be encouraging for the development of EBP which necessitates the involvement of OGD (Other Government Departments) to ensure the alignment of all the instruments of government power (Diplomatic, Information, Military & Economic).

β – Beta / Bias

The β measure of decision criteria identified a number of significant differences between the functional groups but none between the different sites. These data suggested that the criteria adopted by the Command Team and the CIACG was significantly more conservative than the rest of the CTFHQ. These 2 groups were less likely to accept false information as true but as a consequence they were more likely to correctly reject false information. In this situation, neither bias (the CIACG/Command team vs rest of CTFHQ) is better than the other but it **may** provide a method of identifying those individuals who have a better understanding of the situation.

Both the Command Team and CIACG had a better understanding of the Commander's intent and they also had a different criteria (β) to the rest of the CTFHQ. This is an interesting finding that was not anticipated.

Possible explanations of this finding could include the Command Team and the CIACG being knowledgeable about the Commander's intent and were confident in their understanding, as shown by the confidence data, and therefore were more ready to reject some information about the Commander's intent as false.

Alternatively, the members of the CIACG and the Command Team were more experienced and older personnel and may adopt a more conservative bias than the rest of the CTFHQ and it is this artefact that is correlated with greater d'.

However, if this was not the case it may be possible to use criteria (β) to assist identifying those individuals who have a better understanding of a situation when there is no obvious benefit from adopting a particular criteria (i.e. there are no stringent effects of convicting an innocent man or setting a guilty woman free). More data about this situation would need collecting and a proper explanation of the route by which different β would bestow higher sensitivity is required.

This issue was further investigated with the analysis of the Probability of false alarms and hits from the raw data (see Figure 11: Mean Probability of False Alarms by Functional Group and Figure 12: Mean Probability of Hits by Functional Group). Analysis of the data suggested that there was no difference between the functional groups in their ability to "hit" true probes. However, the Command Team and the CIACG were more able to identify false information. This suggests that there is some special quality of these 2 groups.

Assessment of the general approach to quantitatively measuring Commander's intent

Hypothesis 1 stated that using probes and SDT analysis would provide clear insight into the promulgation of commander's intent. The findings discussed above tentatively support this hypothesis. For example, one would expect the Command Team to have a greater level of understanding of commander's intent and this finding corresponds with their confidence ratings. Likewise it is understandable, if not predicted, how the CIACG and Plans, could have a good understanding of intent.

A number of additional factors would appear to support the usefulness of the approach. The completion rate of the questionnaire was very high (91%) and the data would appear to reflect planned and unplanned observations during the event (Farrell, 2004). Farrell (2004) concluded that the level of common intent was low due to the large number of interventions by the Command Team needed to keep the planning process on track.

Conclusion

This paper describes the method and results used by NITEworks to quantitatively measure the promulgation of commander's intent in a simulated CTFHQ during MNE3 across 6 different sites.

It is possible to draw the following conclusions:

1. Overall it would appear that the Commander's intent was not dispersed beyond the Command Team, CIACG and Plans. This may be of concern to policy makers since NEC is credited (House of Commons defence Select Committee, 2003) with the potential to increase the understanding of commander's intent and that there is a requirement for an improved understanding of the Commander's intent for Effects Based Operations. These findings should be tempered with the caveat that these data represent only the very first steps in trying to quantitatively measure the Commander's intent and they were collected in an experimental environment using a coalition type command structure.
2. The method described in this paper appears to measure the promulgation of commander's intent across different sites. The method appears to capture sensible findings, such as the command team has a better understanding of the command intent than the rest of the CTFHQ. Collecting good data in a "field experiment" is extremely difficult and this method appears to be able to collect robust data without undue interference with the HQ staff.
3. It would not be sensible to analyse difference across the different dates/events regarding the modifications of command intent. It could be the result of differences in the situation or it could be attributed to changes in the probe/target difficulty.
4. It would appear that the promulgation of commander's intent is influenced by organisational proximity to the Commander and not geographic proximity to the Commander. This is supported by the US site not having a greater understanding of commander's intent than other nations but those in the Command Team (organisationally and geographically close to the Commander) did have a good understanding of commander's intent. The Plans team, who were geographically remote but organisational close to the Commander also had a reasonable understanding of command intent.
5. Military experience would not appear to be essential to understanding and working with commander's intent. The CIACG, who were not composed military servicemen, had a high d' (sensitivity) which was significantly better than some of the military personnel. This may be a very encouraging finding for the development of EBP where the engagement of non-military experts with the military planning process is a precursor for the alignment of the instruments of power.
6. An extremely experienced subject matter expert considered the level of understanding of the Commander's intent throughout the CTFHQ as very low. In his personal opinion he would not settle for performance below 90% from any of his staff. However, this is the first reported instance of quantitatively measuring commander's intent and it is not known whether this is a normal, even though it may not be desirable.

References

Alberts and Hayes, R. (2003). Power to the Edge. CCRP, Publication Series.

Condell, B. and Zabecki, D. (2001) On the Art of German War (Translation of TruppenFuhrung parts I & II, 1933), Lynne Rienner Publisher, Boulder and London.

Farrell, P. (2004) Measuring Common Intent during Effects Based Planning. CCRP, San Diego.

McGuinness, B. (2004) Quantitative Analysis of Situational Awareness(QUASA):Applying Signal Detection Theory to True/False Probes and Self-Ratings. CCRP, San Diego.

House of Commons Defence Committee (2003) A New Chapter to the Strategic Defence Review. Sixth report of the session 2002-2003 order to be printed 7th May, 2003.

JWP 0-01.1 British Defence Doctrine

Keegan, J. (1999) The Mask of Command: A study of Generalship. Pimlico, London.

Pigeau and McCann (2000) The Human in Command: Exploring the Modern Military Experience, edite by McCann and Pigeau, Kluwer Academic/Plenum Publisher, New Year, 2000. Chapter 12: Redefining Command and Control. Pp165-184.

Sun Izu, (2003) The Art of War. John Minford translation. Penguin, London.

Wickens, C. D. (2002) Elementary Signal Detection Theory. Oxford: Oxford University Press.