THE INTERNATIONAL

C2 JOURNAL

VOLUME 3, NUMBER 2, 2009

Exploring Alternative Edge versus Hierarchy C2 Organizations using the ELICIT Platform with Configurable Chat System

Peter Thunholm
Ng, Ee Chong
Mervyn Cheah
Tan, Kin Yong
Nency Chua
Chua, Ching Lian

CCRP
Focus
& Convergence
for Complex Endeavors

THE INTERNATIONAL C2 JOURNAL

About the Journal

The International C2 Journal was created in 2006 at the urging of an international group of command and control professionals including individuals from academia, industry, government, and the military. The Command and Control Research Program (CCRP, of the U.S. Office of the Assistant Secretary of Defense for Networks and Information Integration, or OASD-NII) responded to this need by bringing together interested professionals to shape the purpose and guide the execution of such a journal. Today, the Journal is overseen by an Editorial Board comprising representatives from many nations.

Opinions, conclusions, and recommendations expressed or implied within are solely those of the authors. They do not necessarily represent the views of the Department of Defense, or any other U.S. Government agency.

For more information

Visit us online at: www.dodccrp.org
Contact our staff at: publications@dodccrp.org

Exploring Alternative Edge versus Hierarchy C2 Organizations using the ELICIT Platform with Configurable Chat System

Dr./LTC Peter Thunholm (National Defence College, SE)

Mr. Ng, Ee Chong (DSO National Laboratories, SG)

LTC (NS) Mervyn Cheah (Singapore Technologies Electronics, SG)

Mr. Tan, Kin Yong and Ms. Nency Chua (Future Systems Directorate, SG)

Mr. Chua, Ching Lian (Defence Science and Technology Agency, SG)

Abstract

The Edge Organization is currently a primary research and experimentation agenda within the C2 research community. By combining the Experiment Laboratory for Investigating Collaboration, Information-sharing, and Trust (ELICIT) Multiplayer Intelligence Game with the use of CHAT, we designed an experiment to explore the effect of different rules of information-sharing, communication, and decision-making on the performance and behavior of three different permutations of Edge versus Hierarchical Organizations. Our main findings suggest that when an intelligence organization is tasked to analyze incoming data and decide on an interpretation of these data, the edge organization outperformed both the traditional hierarchy and the edge-hierarchy hybrid over decision speed, decision

accuracy, and level of shared correct awareness of the threat situation. The hybrid organization and the traditional hierarchy performed equally well on decision accuracy and shared awareness, but the former made decisions faster than the latter. One possible explanation for our results is that in the hierarchical structure, the processing of information takes place in two different levels subsequent to each other in time, but in the Edge there is only one processing level.

Introduction

Network-Centric Operations (NCO; Alberts, Garstka, and Stein 1999; Alberts and Hayes 2003) proposes a shift from the traditional military hierarchical command philosophy to a structure where forces are more nimble and operate on networks to increase their shared awareness as well as to self-synchronize with one another, herein called as the *edge organization*. Alberts, Garstka, and Stein (1999) pointed out that the translation of the NCO concept into a real operational capability requires more than the implementation of information technology and networks. They defined a Mission Capability Package (MCP) comprising concepts of operation, C2 approaches, organisational forms, doctrine, force structure, support services, and the like that is required to leverage information superiority in the realisation of NCW. Also, Alberts and Hayes (2003) maintained that empowerment involves "the ability to provide and access available information and expertise and the elimination of procedural constraints previously needed to deconflict elements of the force in the absence of quality information" (Alberts and Hayes 2003, 5). This paper aims to present the research and experiments carried out jointly by the Singapore Armed Forces Center for Military Experimentation (SCME), the DSO National Laboratories of Singapore (DSO), and the National Defence College of Sweden, towards Edge organization concepts. In the first part of the introduction the authors will elaborate on the concept of C2 in military operations and how it relates to the Edge concept. The paper will then elaborate the mechanics of self-synchronization in relation to

the Edge concept and then introduce the experimental platform that was used for this study. Finally, the paper will present in detail three different C2 concepts that were explored, together with the purpose of the study which put forth the authors' expectations on the outcome of the experiment.

The effectiveness of the military outcome is extremely dependent on, and highly intertwined with, how it applies Command and Control (C2) to the force, in order to achieve its predefined mission objectives through a series of planned actions. There is a chain of effectiveness in what the military does. On the one hand, there is effectiveness of the supporting tasks to the mission objectives such as fires, intelligence gathering and logistics, and on the other, there is the accomplishment of the mission objectives at the tactical, operational theatre and strategic levels. To achieve the above, the military designed a doctrine of fighting and spends a large part of the time training and educating the forces on standard operational procedures, processes, orders, cultures, teamwork and, of course, the art of fighting in combat (e.g., US Army 2003; 2005).

C2 in itself is a complex and dynamic task. Once the mission objectives have been defined, it may only be a matter of hours that the mission objectives and the constraints by which the force is to accomplish the mission would change, as experienced in past conflicts and war (e.g., Simpkin 1985). The planned action becomes outdated and there is a need to dynamically review the mission objectives and constraints continually in order to develop adjustments to the plan of action. The possibility that the adjusted planned action will need to be changed again is high, since each time the "thinking" enemy is also adapting and changing their strategy, objectives and planned actions (e.g., US Army 2003).

Different methods of C2 may be employed to effect the dynamically changing plans. These include centralized command and centralized control; centralized command and decentralized control, decentralized command and centralized control and finally, decentralized

command and decentralized control. Centralized command and control has the advantage of prioritising the allocation of resources to the theatre, but this is seldom practiced by the Army and the Navy of most nations due to the structured echelons of command found in large militaries. At the other extreme of the C2 spectrum is the decentralized command and control concept that allows the force at each echelon to have total command and control, herein also called as the *Edge Organization* (Alberts and Hayes 2003).

For all of the above command and control structures, there are two primary practices to facilitate commanding independent forces with authority, namely mission command and detailed command by planning (Wilbeck 2003). Mission Command seeks to direct the sub-ordinate commander with clear intent such that the lower echelon's planned actions would flow in accordance to the higher echelon's needs and intent (cf., Builder, Bankes and Nordin 2000), while detailed command seeks to exact the detailed plans that each lower echelon would have to abide by, thus only allowing the commander freedom in troop disposition, morale, weapon composition, and perhaps logistics (Wilbeck 2003). The practice differs from military to military. At this juncture, we will not cover the more complex problem of having two or more militaries with different C2 methods or cultures to work together in a coalition setting.

Nevertheless, within a particular military organization, there could be a danger that a particular method of C2 may be the proverbial Achilles heel to the entire operations, whether in loss of lives and resources, time, or the loss of control to gain advantage over the adversary. As an example, a military could practise a centralised command and control of air resources, which for all intents and purposes would be the most flexible way of directing aircraft resources if they are scarce. However, if a request for air strike directed upwards from the Battalion requires concurrence at each chain or level of command, it would take a long lead time for the aircraft to finally

strike on the enemy target, during which the enemy would have had sufficient time to prepare or to take up advantageous positions over our own forces.

There is therefore a need to be adaptive in the nature we command and control our forces. The idea is to adopt a practice for militaries to flexibly move across any one of the described C2 structures in order to adapt to the various scenarios such as maneuver, precision strikes, logistics, etc. over several contexts such as homeland security, operations other than war, and war itself.

As a way forward to adopt this approach, the Singapore Armed Forces Center for Military Experimentation (SCME) began an experiment campaign in 2006 to determine if troops could achieve what is called the *Distributed and Integrated Command Environment* or DICE for short. Cheah and Fong (2006) explained that DICE proposes a command environment that would allow the Commanders or the organization to adopt not only the traditional military hierarchical command philosophy, but also a force structure where the edge elements, that is, the disparate fighting units, are empowered with the information they need as well as the authority to collaborate and self-synchronize in the effective execution of distributed and dynamic operations as they adapt to the changing battlespace situation. Translating, DICE essentially is a concept for forces to work distributed across the battle space, in turn leading to a less hierarchical force structure and allowing a sub-ordinate commander, for instance, to address his issues not only to his immediate superior but also to a circle of experts and higher command through a collaborative operational picture. The US Marine Corps concept of distributed operations is a similar effort to DICE (Hanlon 2004; Schmidle and Hoffman 2004), but without using the word "self-synchronization" at this juncture.

SCME conducted a limited objective experiment (LOE) in an air-land scenario and consisting of forces from both the Air Force and the Army, showed successful results in the adoption of the edge structure where authority for helicopter troop carrying operations

(the MCP) lie with all the heli-pilots and the respective ground commanders going on-board the craft (Cheah and Fong 2006). However, such experiments conducted by SCME, while realistic in operational approaches is also quite costly. To fulfill and experiment all the various MCP that can adopt the DICE concept, without an initial laboratory test for each one, could turn out to be a very expensive affair, and may lead to the experiment controllers being more risk adverse in their experiment approach.

The Effectiveness of the Edge Organization – Empirical Findings

No reports have been found in current C2 research literature of successful attempts to operationally implement an Edge approach. There is, however, some support that self-synchronization (which is a vital part of the Edge concept) can be effective in some C2 situations and some of that research will be presented here. Brehmer (1998; 2009) used a micro-world simulation of fire-fighting to study distributed decision making and found that a network without a central node (a commander giving orders), where all the local commanders could communicate both locations of their own units and their intentions, could outperform a hierarchical network under circumstances of time-pressure.

Dekker (2006) used other types of computer simulations to address specifically the question of effectiveness of self-synchronization under different circumstances, involving network capacity, problem complexity and time-pressure. Dekker compared three different types of networks; one self-synchronized, without a command node, with all nodes (local commanders) on the same level; one centralized, with one command node and all other nodes as subordinated commanders; and one hierarchical network, with three levels of command. His main conclusion was that in complex tasks with time-pressure, and a high capacity network, the self-synchronized network outperformed the centralized network. In simple tasks, without time-pressure the

centralized network performed better, mainly because in the centralized network an optimum solution could be found. However, in real life, especially on the battlefield, the expectation to find an optimum solution is considered highly unlikely. Dekker also found that the hierarchical network, which represents a traditional military C2 structure, could be made to work more effective if the lowest nodes (local commanders) were allowed and able to self-synchronize.

Another series of experiments on self-synchronization using a micro-world of fire fighting was conducted by Rigas, Person, and Brehmer (2005). They tested self-synchronization as a command structure compared to a traditional hierarchical structure and they also compared two different kinds of unit structures in the fire-fighting units. Their main conclusions in those two areas after five different experiments were that self-synchronization was the more effective command structure under circumstances of strong time-pressure and high complexity given that (1) intentions could be presented graphically, (2) the need for synchronization was high, (3) the command intent was obvious, and (4) all fire-fighting units had the same (divisional) structure. Concerning the structure of the units it was shown that the division structure (when all local commanders controlled units containing all types of available assets) was more suited for self-synchronization than a force structure based on a functional structure. A functional structure is when command units consist only of a specific type of asset. These studies on self-synchronization using micro-worlds and the computer-based simulations that were reviewed indicate that (1) time-pressure, (2) the need for coordination, (3) task complexity, (4) network capacity, and (5) the ability for a higher Commander to create an unambiguous command intent for subordinate units to self-synchronization around, all are important factors influencing the effectiveness of self-synchronization versus some degree of centralization.

The main conclusion to be drawn from all the experiments taken together is that these results support the idea that increasing the ability for local commanders to self-synchronize should be beneficial for

mission-effectiveness in some scenarios, at least if local commanders are aware of locations and intent of the other units as well as of the unambiguous intent of the higher commander, if the network has enough capacity and if a fast response is required.

The ELICIT Multiplayer Intelligence Game

Other efforts in experimenting with edge organization have been conducted by the CCRP community such as Evidence Based Research (Martin and McEver 2008), the Naval Postgraduate School (Lewelling and Nissen 2007), and Parity Communications (2006), the company that developed a platform called the Experiment Laboratory for Investigating Collaboration, Information-sharing and Trust or ELICIT for short. ELICIT is an initiative sponsored by the CCRP with the initial purpose (for the C2 research community) to research and experiment differences between hierarchical and edge organization concepts. The original ELICIT game was designed to compare the edge versus the hierarchy, therefore the independent variable in ELICIT was whether a team is organized using traditional Hierarchical Organization or Edge Organization principles. (The current version of ELICIT is more flexible and allows for testing of more type organizations; see Ruddy and Nissen 2008).

ELICIT requires a team of 17 subjects, randomly assigned with pseudo-names, performing the roles of intelligence analysts to collaborate, in a network centric, information processing environment, with the goal to identify a fictitious and stylized terrorist plot. The players receive a number of data messages in a fixed format (called *factoids*) and can communicate with other players by sending factoids, (share), or by posting factoids on a web page, (post) where other players can read them (pull). In the ELICIT game, the experimental task is for every subject to identify the *who* (a group), *what* (a target), *where* (a country), and *when* (a date and time; actually two identifications) of an adversary attack based on the *factoids* that become known to the team members during the game. Putting the game in real-world

context, the organization can be seen as an intelligence organization that has to analyze incoming data and inform its client (or government) about the assessment.

Enhancing Relevance to Real-World C2 Organizations

Similar to previous experiments performed under laboratory conditions, the ELICIT Multiplayer Intelligence Game formed the basic framework for our experiment and we used software version 1.0.0.6 of ELICIT (Ruddy 2007). However, the ELICIT software version available to us included no means of communication among the players, of emergent ideas or conclusions related to the task they were solving, except for sending and receiving fixed-format factoids. In other words, the players were not able to communicate other information besides just factoids, and we regarded this limitation to be a constraint to the ability to self-synchronize (cf., Duncan and Jobidon 2008). In order to enhance the relevance of our tested concepts to real-world organizations, but without altering the features of the ELICIT software, we operated CHAT (free text messages in different configurations) as a communication channel among the members in the organization to get a richer and more realistic game environment in terms of communication capabilities. There are at least two good reasons for using text CHAT instead of some other type of communication channel. The first reason is that CHAT has become an increasingly popular C2 communication tool within many military services and branches (Berube, Hitzeman, Holland, Anapol, and Moore 2007; Eovito 2006; Simpson 2006). Although no formal evaluation efforts to compare the effectiveness of text-chat to other means of communication has been found in the literature, Eovito (2006) reported a dramatic increase in the use of CHAT in all types of field operations since 2002 and that CHAT is now a common channel for communication in most C2 structures. The second reason is that CHAT is fast and responsive enough for the purpose of carrying out group discussions and the discussions are all logged down in textual format. The logged discussions allow for

useful post experiment analysis (as demonstrated in this experiment) without having to transcribe any verbal conversations. The details of the CHAT configuration are described in Appendix I under the heading "Briefing Instructions." The anonymity of each player was secured by using the same names in the CHAT configuration as in the ELICIT game set-up.

Objectives of the Experiment

As presented in the introduction, there is reason to assume that some of the problems facing commanders carrying out new kinds of operations (e.g., stability operations) could be attributed to the maintenance of a less than optimal organizational structure of the force (i.e., a strict hierarchy; Dynes 1994). At least in theory, and based on computational and laboratory tests (as reviewed in the previous parts of the introduction) the edge-organization seems to be better suited to deal with these problems, and that is why efforts to further investigate the potential advantages and disadvantages of the edge versus the hierarchy is important. From a C2 research perspective we thought that a reasonable purpose when exploring different C2 organizations would be to explore which type organization concept best contributes to "increased mission effectiveness" because how to accomplish increased mission effectiveness is one of the main interests in much of the NCO literature (e.g. Alberts and Hayes 2006a; 2006b). Based on the same literature, we also identified that both (decision-) speed and accuracy should be important aspects of mission effectiveness.

Thus, the purpose of this experiment was to explore and compare the effectiveness of the edge organization versus a traditional C2 hierarchy as well as a hybrid of the two C2 organization, where the strict rules of the hierarchy are relaxed in some important aspects (i.e. concerning access to information, communication across organizational boundaries, and decision making), but to all intents and purposes, the organization is still a hierarchy. There are several rea-

sons for putting a hybrid C2 organization to test. For example, in some missions such as peace-enforcement, humanitarian aids, emergency response and others, it is not viable to get a sole commander on the ground to make the decisions as there are many stake holders beyond the military influence. In other situations, there are often several coalition partners or several co-operating organizations that have a stake in decisions to be made as a committee (consensus) or at least voice the decision to be taken by majority or plurality. Also, access to information is often enough less stove-piped and restricted as it used to be, because information can now be accessed through common websites. A third reason for testing a hybrid C2 structure is that email and chat has become quite prominent in recent years and has made communication across organizational boundaries easier than before. We wanted to compare with respect to a typical organization that make sense not only theoretically but also from a practitioner's view regarding access to (and sharing of) information, communication, and organizational decision-making. A fourth reason to explore not only a traditional hierarchy versus an edge can be expressed in terms of the NATO NEC C2 Maturity Model (NATO SAS-065 Research Task Group, 2009). According to this model a Traditional Hierarchy represents *De-Conflicted C2* which puts the Traditional Hierarchy in a region close to one of the corners of the three dimensional *C2 Approach Space* (the least mature C2 approach). The Edge organization (as we define it here) on the other hand represents *Edge C2*, which occupies a region in the opposite corner of the C2 Approach Space. By adding the *Hybrid*, we cover a region in between the other two in the C2 Approach Space; the Hybrid should be somewhere in the area of *Coordinated C2* according to the model.

Different Methods of Making Decisions

One aspect that we wanted to include in our investigation was organizational decision making. A traditional hierarchy normally makes decisions by the commander. In the edge organization it is not so

clear how the organization as such makes decisions, but the general idea seems to be that in the edge, decision rights may be somewhat distributed throughout the actors in the organization (Alberts and Hayes 2003). There are however other ways for an organization to make a decision. For example, making decisions in a committee or coalition, where the decision makers all represent different areas of responsibility or expertise and they are forced to make a decision on consensus, or at least, by majority. Another example is if an edge organization has to come up with a (common) decision. Then such a decision could be made by majority or plurality or some other decision rule. We set out to combine both different configurations of CHAT and different decision rules with the basic ELICIT game in order to explore and compare three different kinds of organizations and they will be introduced briefly in the following.

Traditional Hierarchy

The first structure we decided to explore can be seen as a *Traditional Hierarchy (TH)*, (see also Appendix I). It has four functional or specialist groups of *analysts,* and each group is directed mainly towards one of the four areas, posting information they receive individually on a team website (the where-team has a where-website, etc.). They are the only analysts who have access to that website. Information they receive that does not concern their specific area can be sent to specialists in other teams (but not posted on their team websites). Heading each of the four teams is a *team leader,* who has the same access to information as the members of his team. On top of the four team leaders is the *cross-team coordinator* who is the head of the organization in the ELICIT game. The cross-team coordinator has access to all functional websites and that person, as well as the team leaders, also participates in the analyst work, receiving information on the terrorist plot through the system. In this organization the cross team coordinator has a key role concerning the completeness

of situational awareness within the whole organization, because the cross-team coordinator is the only one who has access to all four web sites.

Regarding communication through CHAT, in *TH* all analysts can communicate with each other, also across teams, and with their own team leader, but they cannot communicate directly with the commander two levels up, the cross-team coordinator. Such communication must go through their team leader. We regard this restriction as a common trait and process of a traditional hierarchy. All team leaders can communicate with each other and with the cross team coordinator. The ability to communicate through CHAT between analysts of different teams symbolizes what is actually possible and frequently occurring in real-life traditional hierarchies today (cf., Alberts and Hayes 2003, 216-217), and it does not change the formal responsibilities or work flows of the *TH*, but it makes it more operationally realistic.

The organizational decision making in *TH* is done by the cross-team coordinator, who submits the final assessment of the threat (all four parts) when he/she feels certain about it. It is expected that the team leaders for the four different functional teams are supposed to provide the cross-team coordinator with an answer on their specific question.

Edge – Hierarchical Hybrid

The second structure we explored was a *Hybrid (HY)* between a traditional hierarchy and an edge organization (see also Appendix I). It is still specialized with four functional groups of analysts, but here the analysts and the team leaders, have access to all four functional websites. As in the *TH*, all members of the organization will receive individual information vital to assess some part of the four areas. Concerning communication in CHAT, the CHAT set-up in

HY allows everyone to interact with everyone else, so the restriction in the *TH* for analysts to communicate directly only one level up is removed in the *HY*.

Organizational decision making in *HY* is made by a majority in the group of team leaders and the cross team coordinator. When three out of five in the group agree on all four areas, they can submit their common answer as the assessment made by their organization to their client. This way of making decisions actually makes the organization a *two-level hierarchy* and not a three-level (as the *TH*). The cross team-coordinator is not superior to the team-leaders but on the same level in this respect. His/her "vote" does not have more merit than the votes of the other four members of the command group. However, the cross-team coordinator is still formally the head (organizer) of the work and the leader or chairman of the team leaders, although he/she cannot make a formal decision on the threat assessment without the support of at least two of the team leaders (and can of course be out-voted if three of the team leaders agree on a different assessment).

Edge

In comparision to the previous two structures mentioned—the third structure we explored was the *Edge (E)*, (see also Appendix I). In this structure there are no functional (specialist) groups and no team leaders. All members are analysts with the same mission and they are free to choose (self-synchronize) what area they want to focus on. As in the other two organization types, all members of the organization will receive individual information vital to assess some part of the four areas. Concerning communication, the CHAT set-up allows everyone to interact with everyone else.

Organizational decision making is done by majority among the analysts. When nine out of seventeen analysts agree on all assessments, they report or submit their assessment to their client. This way of

making decisions in the edge organization can be seen as a *violation* on the basic idea of the edge or as taking *the edge of the edge.* The basic idea in the theory about the edge organization as it is presented by Alberts and Hayes (2003) is that in the edge, each agent is so well informed on the overall intent and has the same high level of shared awareness as the other agents in order for them to decide individually. Imposing a majority rule for organizational decision making upon the edge organization would be to "miss the point."Although we agree on the basic assumption of the edge, we still wanted to test if the edge could also be effective in a situation where it has to agree on a common decision. In today's world of coalitions and consensus building before action can be taken we thought that imposing a majority decision rule upon an edge organization was easily defendable and also interesting enough to investigate further.

Expectations

Regarding what expectations to have concerning the effectiveness of the different organizational structures, some general guidelines were provided (by the research on self-synchronization previously presented in the introduction) that seems to imply that the Edge organization should be more effective than different types of hierarchies under certain circumstances. This research is however not directly transferable to the ELICIT game situation, thus making it difficult to use those results or research as a base for formal hypotheses.

There is also a large body of research on team-work effectiveness (e.g., Salas, Sims, and Burke 2005; Marks, Mathieu, and Zaccaro 2001) and team-sensemaking (Jensen 2008) that points out the importance of a clear structure and team leadership in order to organize and coordinate the team effort and to be effective as a team. This research could be used as a base for expecting a hierarchical structure with its clear leader appointment and team structure as more advantageous that the Edge structure; however as the ELICIT game was set up, there is nothing preventing an Edge team from

self-organizing, and for leadership to emerge, thus making it difficult to develop expectations based on that research. Earlier research based on the ELICIT game (e.g., Leweling and Nissen 2007; Parity Communications 2006) could also be used as a base for developing expectations. However, earlier research and experiments using the ELICIT did not include both CHAT and the decision-making rules. Thus, adding these features to the type organization concepts tested in this study made it difficult to develop strict hypotheses. Taken together with the combination of the small number of participating teams, resulting in a small basis for proper statistical analysis of results, it restrained us from formulating and testing strict hypotheses. Instead we formulated a few expectations to evaluate and they will be elaborated in the following section.

Accuracy of Decisions Made

Concerning the level of *accuracy of the decision* (threat assessment) delivered by the different organizational types, earlier research gives little guidelines because no decision rules like the ones we used were imposed. We postulated that the decisions made by majority, as in the Edge *(E)* but also, in smaller scale, in the *HY,* should at least in theory be more elaborately discussed and assessed. The risk that a false conclusion should go undetected should be less, because more individuals with shared responsibility for the correctness of the complete assessment should have processed the information underlying the assessment. Thus we expected the *E* (which should involve the most number of players debating and agreeing on the complete answer) to deliver the highest level of accuracy in the decision, followed by the *HY* and then *TH*.

Decision Making Time

Concerning the *length of time* it should take for the different types of organizations to come up with their decision (assessment), earlier research indicated shorter time for the Edge (Leweling and Nissen 2007). However, as their experiment did not include the majority rule for decision making, we postulated the implementation of this decision rule into the edge organization to be rather time consuming for the members of the edge to derive a majority agreed decision as compared to the situation where a single decision maker makes the decision, as in *TH*, or when a smaller group has to agree on the decision, as in *HY*. Thus our expectation was that the *E* should need more time to submit its organizational decision on a full assessment of the terrorist threat, followed by the *HY* and then the *TH* as the fastest organization to make decisions.

Level of Shared Awareness

Regarding the level of *shared awareness* of the threat we defined it as the level of common agreement among the players in an organization, on the correct assessment in all four areas. The interesting point to note here is whether a certain organization structure would result in a higher level of *correct* shared awareness or less. Shared *incorrect* awareness, if formed, is not a good thing for any organization as it means that the organization commonly believed in a wrong perspective of the truth. The reason for including such a measurement is that shared awareness is a part of the NCO value chain, were it is described as a precondition for successful self-synchronization (Alberts and Hayes 2003). Our definition of shared awareness means that this measurement in part captures the same phenomenon as the measurement of decision accuracy, but here we are interested in the overall proportion of players in each organizational structure that finally agrees on the correct assessment. On this, earlier results gave mixed signals because in the study by Leweling and Nissen (2007), and by Duncan and Jobidon (2008) there were no differences in level

of shared awareness between the edge and the hierarchy, but in the study by Parity Communications (2006), the edge performed better than the hierarchy. We postulate that the edge, at least in theory, should demand more individuals to engage in constructive discussions regarding the full assessment of the game outcome as opposed to the situations in *TH* and *HY* where analysts assigned to solve parts of the task would only discuss relevant components of the assessment. Therefore we expect the level of shared awareness to be highest in the *E*, followed by the *HY* and then the *TH*.

Method

Design

Although we did not design this study to test hypotheses we still employed an experimental design. The independent variable was the *type of organization*, in three different levels: (1) *Traditional Hierarchy (TH)*, (2) *Hybrid*, between a traditional hierarchy and edge organization *(HY)*, and (3) *Edge (E)*. The difference between the three types is described in the previous section and in Appendix I.

Participants

The experiment involved the entire graduating cohort of senior male and female military officers (Major, and some Lieutenant Colonel in rank) from the Singapore Command and Staff College (SCSC). The participants were organized into seven teams, with a random mix of training backgrounds (army, navy, and airforce), and they were all reasonably equal with respect to their military experience and age. Since the team compositions were comparatively on par in terms of operational capability and seniority, we did not measure any back-

ground variables, and also partly because earlier testing did not show any substantial relationship between experience from analyst work and performance on the ELICIT game (Leweling and Nissen 2007).

Procedure

Each participant went through one formal and one applied training run on ELICIT before their actual run. The total time for training was about 1.5 hours. One scenario was used to conduct the training runs, and another scenario was used to conduct the actual runs. Before the commencement of each run, the teams were briefed based a standard set of instructions developed. The detailed procedure is described in Appendix I. The breakdown of the C2 concepts experimented by all 7 teams are listed in Table 1.

Table 1. Number of Teams in Each Experimental Condition

C2 Concept	Numbers of Teams
Level I - Traditional Hierarchy (TH)	2
Level II – Hybrid HY)	2
Level III – Edge (E)	3

Dependent Variables

Dependent variables were (a) *organizational decision accuracy*, inferred (1) from the proportion of correct answers on all four areas (i.e., requiring one answer regarding *who* is behind the plot, one answer regarding *what* is the target; one answer regarding *where* the target is located and two answers regarding *when*, date and time, it will take place) delivered by the head of the organization, for that specific

organizational structure; and (2) from the proportion of correctly identified plot details among all members of the organization (if all members correctly identify all five plot details, the maximum score is 85); (b) *decision speed*, inferred from the time taken for the organizational decision maker to deliver the team's decision or assessment on all four areas, and we measured the variation in time between the first member in each team to submit his/her final answer and the last one to do so; (c) the level of *correct shared awareness* obtained among the members of an organization, inferred from the proportion of correct answers given on all five questions by the members of an organization and (d) the *working process*, inferred as a qualitative assessment on (1) how the information was accessed and shared, and (2) patterns of communication through CHAT, and decision making within the organization.

Measurements

As in prior experiments based on the ELICIT software, most of the dependent variables were computed from the Elicit Transaction Log files. The computation taken for each dependent variable is summarized in Table 2.

Table 2. Computations Made for Each of the Dependent Variables

	Experimented C2 Concept		
Dependent Variable	**Traditional Hierarchy (TH)**	**Hybrid (HY)**	**Edge (E)**
Organizational Decision Accuracy	(1) Solution posed by Cross Team Leader	(1) Solution arrived by consensus (3 or more agree) between leaders	(1) Solution arrived by consensus (9 or more agree)
	(2) Proportion of correctly identified plot details by all members	(2) Proportion of correctly identified plot details by all members	(2) Proportion of correctly identified plot details by all members
Decision Speed	(1) Time taken for Cross Team Leader to decide	(1) Time taken for 3 or more among the leader group to agree on solution	(1) Time taken for 9 or more in the whole organization to agree on solution
	(2) Variation in time between first and last member to submit a full answer	(2) Variation in time between first and last member to submit a full answer	(2) Variation in time between first and last member to submit a full answer
Level of Correct Shared Awareness	Proportion of Org with 100% correct answers	Proportion of Org with 100% correct answers	Proportion of Org with 100% correct answers

In addition, context of the detailed CHAT activities by each team during the experiment can be obtained from the Log files of the Chat-rooms. Rather than a descriptive analysis of the contents of all the CHAT activities, we approached the analysis of the CHAT activities by categorizing the CHAT messages sent by each participant into 5 main types, namely *Posts*, *Analysis*, *Coordination*, *Answers,* and *Others*. *Posts* refers to CHAT activities during which the participants ***reproduce*** the information given by the factoids in ELICIT into the CHAT messages, ***without*** any interpretation or analysis done. *Analysis* refers to CHAT activities during which the participants consolidated the information from different factoids, created their own interpretation, shared their interpretation, posed questions to clarify other participants' interpretations, and asked for more information. *Coordination* refers to CHAT activities during which participants issued instructions to other participants to either collect information, perform some form of analysis, or to form focus groups to solve specific parts of the problem. *Answers* refers to CHAT activities during

which the participants provided their proposal, in parts or whole, of the final answer. All other activities, including non-ELICIT related CHAT activities such as "Hi!," are categorized under *Others*.

When all the messages in the CHAT logs of each run were categorized, analysis of the CHAT activities were performed by comparing across all the concepts, the overall volume of CHAT messages, defined as the overall CHAT traffic, and the volume of each type of CHAT activities under *Posts*, *Analysis*, *Coordination*, and *Answers*. We considered *Others* as noise CHAT traffic and would not interpret these specifically.

Lastly, analysts were stationed to monitor and note observations for each team during the conduct of the experiments.

Statistical Analysis

Statistical analyses were performed through STATISTICA 7.0. (www.statsoft.com). All variables were checked for normal distribution and the correlation (Pearson *r*) between them were computed. We also computed mean values and standard deviations for the outcome of the five different measurements on the three dependent variables, organizational decision accuracy, decision speed, and level of shared awareness for each condition (*TH, HY,* and *E*). We did not apply significance testing of the differences between mean values because there were too few observations (two teams) in each condition. Instead, the results were evaluated by means of Cohen's *d*, a measure of effect size (Cohen 1972; 1992). Cohen's *d* is not a significance test. Cohen's *d* expresses the size of the effect as a proportion of the standard deviation from the mean value. Thus a Cohen's *d* of 0.50 means that the effect of the experimental variable was half a standard deviation, and a Cohen's *d* of 1.0 means that the effect was one standard deviation (quite unusual in behavioral science experiments). In view of the fact that this study was not a formal hypothesis testing study, this form of analysis should be sufficient.

However, because of few observations in each variable there is uncertainty regarding the true value of the standard deviations, and the computation of the effect size, thus these statistics should be interpreted with some caution.

Results

Dependent Variables

The results for the dependent variables are summarized in Table 3.

Table 3. Individual Team Results on Three of the Dependent Variables

C2 Concept	Organizational Decision Accuracy 1 (Correct / Incorrect answer by the Head of the org)	Organizational Decision Accuracy 2 (Correctly identified plot details out of 85 among all members: number/ proportion)	Decision Speed 1 (Total time, minutes)	Decision Speed 2 (Variation in decision speed between fastest and slowest org member in minutes)	Shared Awareness (Number of Org members/ Proportion of Org with 100% correct answers out of 17)
Hierarchy1	Incorrect	62/73%	41	10	4/23.5%
Hierarchy2	Correct	33/39%	44	20	1/5.9%
Edge1	Correct	80/95%	36	1	16/94.1%
Edge2	Incorrect	48/56%	35	7	0/0%
Edge3	Correct	82/96%	25	6	15/88.2%
Hybrid1	Incorrect	44/51%	36	3	0/0%
Hybrid2	Correct	68/80%	41	4	9/52.9

Outliers

One of the teams in the Edge condition (indicated as a shaded entry) did not follow the instructions completely. In the after-action review performed after the training run, the team members decided that they preferred to have some structure in their Edge run so they appointed the same player who had acted as cross-team coordinator in the training run to perform a similar function also in the edge run. The appointment of a leader *before* the game had started meant that the team had removed a major part of the collective responsibility as an inherent Edge trait as well as the ability for emergent control behavior to occur. The team was therefore not operating as an Edge organization the way we intended. We decided to treat that team as an outlier, but we still use this team for some of our qualitative evaluations later in the text, but it was not included in the quantitative analyses for each dependent variable.

Correlation Between Outcome Variables

The outcome variable *decision accuracy* is bivariate (*correct* answer on the four questions or *incorrect* answer) and thus not normally distributed. Evaluation of assumptions of normality regarding the outcome variables *decision speed* and *shared awareness* showed that these variables were reasonably normally distributed with skewness and kurtosis less than +-2. The correlation between decision speed and shared situation awareness was rather substantial but not statistically significant at the 5%-level ($r = -0.50$, $p > .05$).

Organizational Decision Accuracy

In the edge condition both teams made correct majority decisions (i.e., at least nine players agreed on a common assessment on all four questions). In the *HY* and the *TH* one of two teams in each condition made correct decisions. If a correct answer is scored as

2 and a wrong answer is scored as 1, this gives a mean value of 2.0 (*SD* = 0.0) for the *E* and mean values of 1.5 (*SD* = 0.7) for both *TH* and *HY*. The effect size index is 1.01, indicating a large effect; a *d* larger than 0.8 is regarded as a large effect size (Cohen 1977). The result suggests that the teams in the edge condition had higher organizational decision accuracy than did teams in *TH* and *HY*, but the latter two teams performed equally good/poor in this respect. Regarding the measurement on proportion of correctly identified plot details among all members of the organization, the two Edge teams on average correctly identified 81.0 (*SD* = 1.41) or 95% of a maximum possible score of 85. The corresponding figure for the *HY* was *M* = 56.0, (*SD* =16.97) or 66% and for the *TH* it was *M* = 47.5 (*SD* = 20.50) or 56%, a rather substantial difference also reflected by a Cohen's $d > 2.07$ between *E* and both *HY* and *TH*. The difference between *TH* and *HY* was small, reflected by a $d < 0.48$.

Timing and Task Completion

All teams completed the task and submitted an answer although some individual players in some teams did not. For the two teams in the traditional hierarchy condition (*TH*) the average decision time was 42.5 minutes (*SD* = 2.12). For the hybrid between hierarchy and edge organization condition (*HY*) the mean value of decision time was 38.5 minutes (*SD* = 3.53). For the edge organization condition (*E*) the mean value was 30.5 minutes (*SD* = 6.73), and the mean value for all conditions was 37.2 minutes. Thus, the difference between the *TH* and *HY* was 4 minutes; the difference between the *E* and the *HY* was 8 minutes, and the difference between the *TH* and the *E* was 12 minutes which means that the teams in the *TH* condition on average used more than 25% longer time to make a decision than did the teams in the *E*. The difference is substantial and this is reflected in the effect size index: Cohen's *d* = 2.10. Also the differences in decision times between the *E* and the *HY* and between the *TH* and the *HY* are quite substantial as indicated by effect size indexes over 1.3. These results indicate that the fastest organization to make a

decision was the edge, followed by the hybrid and the slowest was the traditional hierarchy. This result is contrary to our expectations, and it suggests that the need to seek a majority consensus that we imposed on the teams in the *E* condition (and also, but to a lesser degree, on the teams in the *HY*) did not slow them substantially. Our results are thus in line with earlier results (Leweling and Nissen 2007) demonstrating that the edge performed faster than the traditional hierarchy. Regarding the variation in decision speed within each org type the Edge teams on average had a mean value of 4.0 minutes ($SD = 4.2$) in difference between the first and the last team members submitting their final answers. The corresponding figures for the *HY* was $M = 3.5$ minutes ($SD = 0.7$) and for the *TH* it was $M = 15$ minutes ($SD = 7.1$). The difference between the *TH* with respect to both *E* and *HY* was substantial as indicated by the effect size index ($d > 1.9$) but the difference between the *HY* and *E* was not. This suggests that in the *TH* there was quite a duration (15 minutes on average) between the first organization member submitting his/her answer and the last one, but in the *E* and the *HY* the time span between the first and the last organization member was far shorter. This also suggests that in the *TH* there was less collaboration and consensus building, leading to each member to work more individually to get a grip of the complete threat situation.

Level of Shared Awareness

In the Edge condition the average number of team members who submitted correct assessments on all five questions (*who*, *what*, *where* and the two *when*) was 15.5 ($SD = 0.7$) or 91%. The corresponding figures for the *HY* was 4.5 ($SD = 6.36$) or 26.4%, and for the *TH* it was 2.5 ($SD = 2.12$) or 14.7%. The effect size index indicates a small effect between the *TH* and the *HY* but a large effect both between the *TH* and *E* (Cohen's $d = 8.23$) and between the *HY* and *E* (Cohen's $d = 2.43$). The result suggests that the teams in the edge had a high level of shared correct awareness of the threat situation, and also considerably higher than did teams in both of the

other conditions. The difference between teams in *TH* and *HY* was unsubstantial, suggesting that these two types of organizations perform equally well/poor regarding the ability among the members of the organization to reach a high level of shared, correct, awareness. All these three main results of the dependent variables are in favor of the edge organization. But before we discuss the reliability and validity of the results we will provide a more detailed description of how the teams in the different conditions actually performed as this will make it easier to understand why some of the differences in the outcome variables explained above occurred.

Comparison of the CHAT Activities and Traffic Volume

Table 4 shows the breakdown of the CHAT activities as well as the overall volume across all seven teams.

Table 4. CHAT Activities for all Teams

C2 Concept	Posts	Analysis	Coordination	Answers	Others	Overall (Total)
Hierarchy1	18	329	120	25	60	**552**
Hierarchy2	6	393	121	32	234	**768**
Edge1	16	82	55	144	94	**391**
Edge2	19	122	71	39	35	**286**
Edge3	3	29	31	57	66	**186**
Hybrid1	34	319	68	40	211	**672**
Hybrid2	12	314	103	73	75	**577**

We observed that the overall volumes of CHAT traffic under the *E* runs were lower than that of *TH* and *HY* runs. Mean value for *E* was 288 (*SD* = 144.9). The corresponding figures for *HY* was 622 (*SD* = 70.7), and for *TH* it was 669 (*SD* = 165.5). The effect size index indicates a small effect between the *TH* and the *HY* (Cohen's $d < 0.5$)

but a large effect both between the *TH* and *E* Cohen's d =2.44) and between the *HY* and *E* (Cohen's d = 2.93). This is counter intuitive to our original expectations as we expected that in the *E* concepts, the CHAT traffic volume will be higher because there is a need to poll everyone in the team to get a consensus on the final answer. The much higher *Analysis* traffic compared to the lower *Posts* traffic indicated that the participants devoted their time in the CHAT to conduct analysis of the information rather than just regurgitating the factoids in the CHAT environment. Lastly, we also observed that having multiple chat rooms did result in duplicated analysis, which meant that questions already raised, discussed and resolved in one CHAT room may also be discussed again by the participants in another CHAT room, due to the segregation of the participants in different CHAT rooms. This result was also clearly evident from the much higher "Analysis" CHAT activities observed in the *TH* and *HY* runs, in which the participants were segregated in different functional CHAT rooms.

Detailed Observations for Traditional Hierarchy Runs

A detailed look at the breakdown of individual answers in the two Hierarchy teams showed that the *Where* teams in both runs had difficulty identifying the answer to the *Where* task (see Table 5). The first team that experimented on the *TH* concept did not manage to give the correct answer for the *Where* part of the solution. It was observed that 2 out of 3 members in the *Where* task got it correct, the *Where* team leader did not accept the answer and hence was unable to provide a correct answer to the Cross Team Leader. Interestingly, we noticed that the second team only managed to give the correct answer because the Cross Team leader worked on the analysis himself and proposed the overall solution instead as he was not able to obtain confirmed solutions from the team leaders.

Table 5. Individual Team Members Results on the Threat Identification Task and the Information-Sharing Activities in the TH Teams

C2 Concept	Info Dissemination: Total SHARE	Info Dissemination: Total POST	Pull from Websites: Total PULL	Team Members (M), Leaders (L), and Cross Team Coordinators (C) who got their task correct							
				Who		What		Where		When	
				M	L/C	M	L/C	M	L/C	M	L/C
Hierarchy1	50	75	238	3	1/1	3	1/1	2	0/0	3	1/1
Hierarchy2	44	80	432	3	0/1	3	0/1	2	0/1	2	1/1

Table 5 also presents the number of *SHARE*, *POST*, and *PULL* of factoids within each team. SHARE and POST is approximately the same number for both, and will be commented further in the following section, but PULL (i.e., reading a factoid from a website) was much more extensive in the second and more successful of the *TH* teams and reflects a more intensive analysis activity in that team.

Using the PAJEK software, network diagrams of the SHARE and POST traffic for each team experimenting with the *TH* concept were obtained. Figure 1 presents the linkages that resulted from the SHARE traffic in both *TH* runs.

In this ELICIT run "Alex" served as *Cross Team Coordinator* and the following names served as Team leaders: *Who* - "Chris," *What* - "Jesse," *Where* - "Pat," and *When* - "Sidney;" this was the same for the Hybrid teams as well, but in the Edge there were only analysts and no appointed leaders.

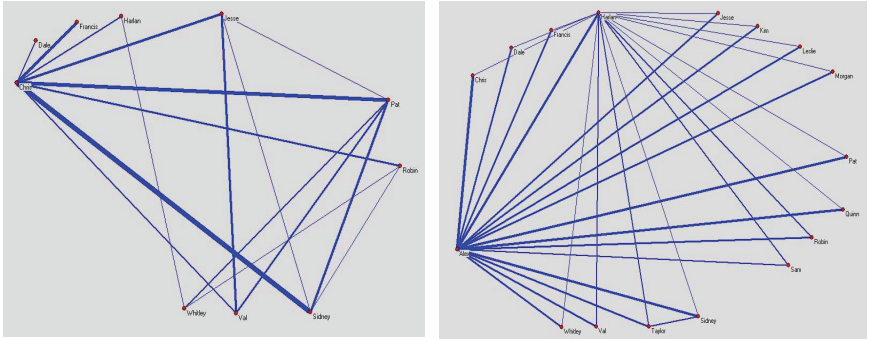


Figure 1. The SHARE Diagrams for both Teams under TH Concepts (left is Hierarchy 1 and right is Hierarchy 2, cf. Table 5)

On the left of the two PAJEK diagrams in Figure 1, only 10 out of 17 participants were visible. This illustrated the lack of active sharing in the Hierarchy 1 Team, as some members of the organization were clearly missing from the network and were hence not sharing factoids with other players in the organization, although the total number of SHARE is approximately the same in the two *TH* teams (cf., Table 5). The thicker lines represented some active sharing between certain members. We therefore observe variance in the participants' propensity to contribute towards common availability of information within the organization. Although sharing was limited to fewer team members in one of the *TH* teams, all critical factoids for identification of the plot detail assigned to the team were available within each team.

The POST network (Figure 2) is similar for both *TH* teams and illustrates the effect of the rule that each team in the Traditional Hierarchy only have access to their own information website.

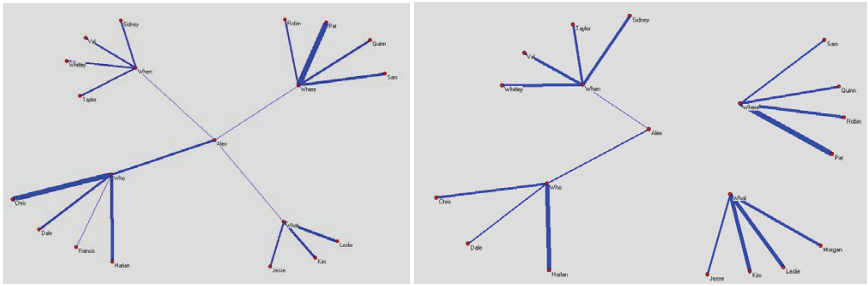


Figure 2. The POST Diagrams for both Teams under TH Concepts (left is Hierarchy 1 and right is Hierarchy 2, cf. Table 5)

CHAT analysis. Under the *TH* concepts, the volume of *Analysis* traffic was much higher compared to the *E* concepts and comparable to the *HY* concepts (cf., Table 4). We attributed this observation to the duplicated analysis that went on in parallel in the functional CHAT rooms, and this conclusion was substantiated by the contents of the CHAT logs for the *TH* runs. By this, as explained earlier, we meant that questions raised, discussed, analyzed and resolved in one CHAT room may also be raised in parallel by participants in another CHAT room. There were also more querying and questioning because of the uncertainty resulted from the incomplete factoids that each participant is receiving from the ELICIT environment. The *Answers* traffic was low for the *TH* runs was expected as the Leaders did not have to poll for consensus before submitting the final answer.

Detailed Observations for Hybrid Runs

A detailed look at the breakdown of individual answers in the two Hybrid runs showed that the two teams experimenting on the Hybrid concept had very different outcomes (see Table 6). A detailed examination of the factoid dissemination showed that one of the members in the Hybrid 1 team did not post a critical factoid on the websites and hence no one else in the organization had access to this piece of

information. Hence all analysis and deductions made by the members and leaders led to the wrong conclusions and the entire organization reached a consensus but on the wrong answer. The Hybrid 2 team managed to achieve good dissemination of the factoids and hence was able to arrive at the correct answer.

Table 6 also presents the number of SHARE, POST, and PULL of factoids within each *HY* team. The PULL activity was more extensive for both of the *HY* teams than for any of the *TH* teams (cf., Table 5) but seems to occur at approximately the same magnitude for both of the *HY* teams, reflecting an intensive analysis activity in both teams.

Table 6. Individual Team Members Results on the Threat Identification Task and the Information-Sharing Activities in the HY Teams

C2 Concept	Info Dissemination: Total SHARE	Info Dissemination: Total POST	Pull from Websites: Total PULL	Team Members (M), Leaders (L), and Cross Team Coordinators (C) who got their task correct							
				Who		What		Where		When	
				M	L/C	M	L/C	M	L/C	M	L/C
Hybrid 1	0	124	554	3	0/1	1	0/0	0	0/0	0	0/0
Hybrid 2	74	125	501	3	1/1	2	1/1	3	1/1	1	1/1

We also discovered that the availability of common websites in the *HY* condition resulted in a substantial increase of the POST traffic compared to the *TH* teams (see Figure 3). The difference compared to the *TH* teams (Figure 2) indicates that all members have access to all websites and they use their ability to post information also on other teams websites. Also, the Hybrid 1 team did not use SHARE at all while the Hybrid 2 team used SHARE even more than the two *TH* teams (see Figure 4), and in a similar pattern as for the *TH* team 2 (to the right side in Figure 1), indicating that the Cross-Team Coordinator was a central node in the share network also in

the Hybrid 2 team. In the Hybrid 1 team however, all participants posted and pulled their factoids from the websites instead of sending (i.e., sharing) them to another individual member.

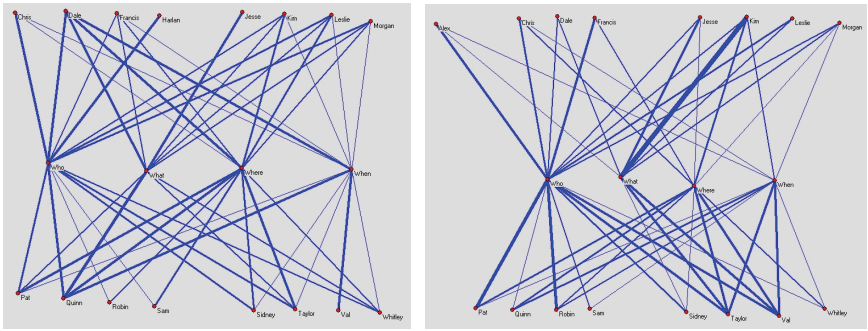


Figure 3. The POST Diagrams for both Teams under HY Concepts (left is Hybrid 1 and right is Hybrid 2, cf., Table 6)

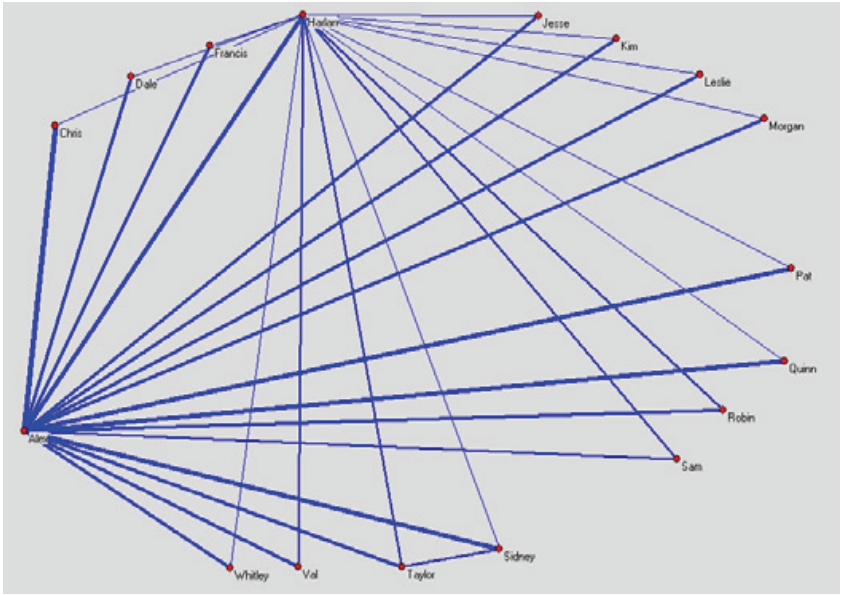


Figure 4. The SHARE Diagrams for one Team under HY Concept (Hybrid 2)

CHAT analysis. For *HY* concepts, we observed a similarly high volume of *Analysis* traffic as in *TH* (cf., Table 4). This is counter-intuitive because the factoid distribution set-up for the *HY* (with common access to websites) in ELICIT is similar to the *E* concepts set-up. However, the results suggested that the *Analysis* traffic is as high as in the *TH* concepts. Again we observed that this high *Analysis* traffic was attributed to the duplicated analysis occurring in the different functional CHAT rooms. The participants discussed their analysis in their respective functional CHAT rooms instead of the uniformed CHAT room. This resulted in an inefficient analysis process which was also presented and explained in the CHAT results for the *TH* concepts. The results also reinforced our belief that although there may be a non-hierarchical set-up in the ELICIT environment, the efficiency at which the analysis of the factoids can be carried out depends on whether the participants performed their discussions in a unified CHAT room or separate functional CHAT rooms. We would discover more on this after presenting the Edge runs results.

Detailed Observations for Edge Runs

A detailed look at the breakdown of individual answers from members of the two Edge teams showed that the two teams experimenting on the Edge concept had similar outcomes (see Table 7). A detailed examination of the factoid dissemination showed that all critical factoids were disseminated and hence everyone in the organization had access to all information. It is also interesting to note that the decision making by consensus leveled up almost everyone in both runs to the complete answer. However, an unexpected observation was made in both runs: After an initial stage of level playing field, some form of leadership emerged as one player in each team took on responsibilities as a team leader (and this is further commented under the CHAT results).

Table 7 also presents the number of SHARE, POST, and PULL of factoids within each Edge team. The PULL activity was lower in Edge 3 than in Edge 1 but both was at approximately at the same magnitude as for the *HY* teams and one of the *TH* teams. (cf., Table 5 and 6), reflecting an intensive analysis activity in both teams. However the SHARE activity was zero in both teams which is similar to the unsuccessful *HY* team, indicating that it was not the inactivity in SHARE per se in the unsuccessful *HY* team that resulted in their failure. The level of POST activity was approximately the same in both *E* teams and approximately at the same magnitude as for the *HY* teams (see Figure 5). This is not surprising because in both *HY* and *E* all team members have access to all websites. The PAJEC diagrams in Figure 5 shows a similar pattern between the two teams, with all members actively posting, but there are individual differences in the level of activity among the team members as reflected by the thickness of the lines in the graph.

Table 7. Individual Team Members Results on the Threat Identification Task and the Information-Sharing Activities in the Edge Teams

C2 Concept	Info Dissemination: Total SHARE	Info Dissemination: Total POST	Pull from Websites: Total PULL	Team Members (M) who got their task correct			
				Who	What	Where	When
				M	M	M	M
Edge 1	0	94	585	16	16	16	16
Edge 3	0	124	410	17	16	17	16

As mentioned earlier, the team that did not perform the experiment according the instructions was identified as an outlier and was not included in this analysis.

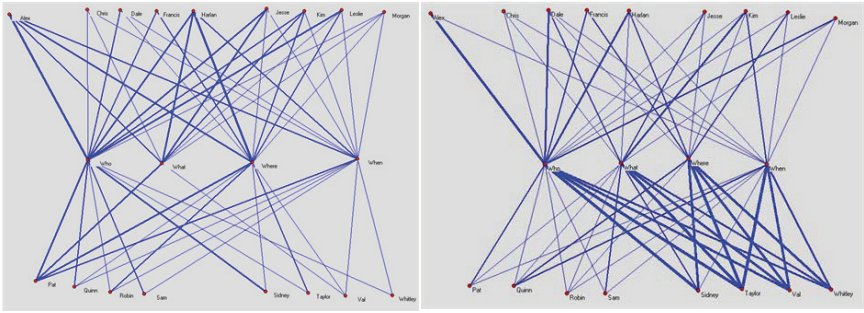


Figure 5. The POST Diagrams for both Teams under E Concepts (Edge 1 is on the left; Edge 3 is on the right, cf. table 7)

CHAT analysis. Evidences from the Chat Log showed that "Kim" for the first team and "Leslie" for the second team emerged as eventual leaders who consolidated the solutions based on consensus building and concluded the runs. For *E* concepts, the results showed that the overall CHAT traffic was low compared to the other two concepts. Particularly the *Analysis* activities were also lower. The main reason for this was that the congregation of everyone in the same chat room reduced the duplication of analysis and hence everyone could see everyone else's suggestions, queries and contributions. The *Answer* activity traffic volume is also surprisingly low, which is counter-intuitive to our expectations as we expected that this activity will actually be higher since there is a need to obtain a larger consensus. However, by looking through the contents of the CHAT logs, we actually observed less challenges being proposed to suggested answers because everyone who had followed the various discussions and analysis seemed to agree on the way the final answer was derived and hence the consensus was actually reached very rapidly in the *E* runs.

Detailed Observations for Outlier Data Point

Apart from the fact the team did not intentionally participate as an Edge organization, the detailed observations also revealed interesting findings. One such observation showed that the team did not perform well partly due to incomplete dissemination of critical factoids, which was due to inactive posting of factoids by certain members of the organization as can be seen from Figure 6, only 14 out of 17 team members did post their factoids, and one of the critical factoids was kept private by one team member. This was also the main factor affecting non-performing run under the *HY* cases.

Table 8. Individual Team Members Results on the Threat Identification Task and the Information-Sharing Activities in the Outlier Edge Team

C2 Concept	Info Dissemination: Total SHARE	Info Dissemination: Total POST	Pull from Websites: Total PULL	Team Members (M) who got their task correct			
				Who M	What M	Where M	When M
Edge 2	0	78	511	13	11	11	0

As can be seen from Table 8, the number of PULL and SHARE activities was approximately at the same level as for the other *E* teams, but the level of POST activities was more similar to the two *TH* teams.

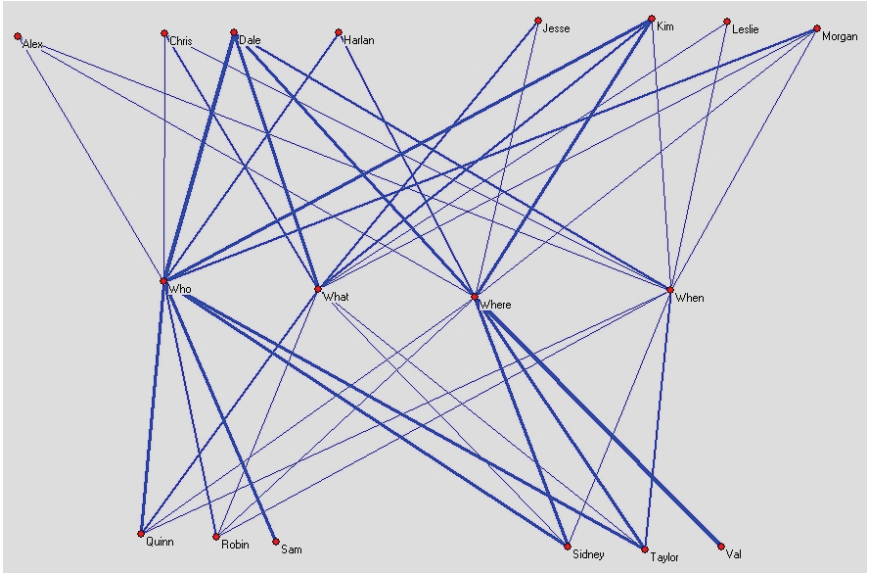


Figure 6. The POST Diagram for the Outlier (Edge 2) Team

CHAT analysis. For this outlier *E* concept run, there was no obvious difference in the observation for CHAT activities compared to the other two *E* concept runs. We gathered that although the team had intended to behave in a non-*E* concept, the benefit of congregating in the same CHAT room also meant they still did not incur duplication of analysis, as was observed in the *TH* and *HY* runs where there were separate CHAT rooms. This actually forms a very interesting data point, in that the volume of chat activities is directly related to the way the CHAT rooms were set-up and may not be as much impacted by the ELICIT set-up or the mindset of the participants.

Discussion

The main findings from this study was that on this task, when an intelligence organization have to analyze incoming data and decide on an interpretation of these data, the Edge organization outper-

formed both the Traditional three-level Hierarchy and the Hybrid Edge/Hierarchy organization on decision speed, decision accuracy and level of shared, correct, awareness of the threat situation. The Hybrid organization made decisions faster than the Traditional Hierarchy, but performed equally well on decision accuracy and shared awareness among the members of the organization. The main reason behind the difference between the Edge and both of the Hierarchies (the *HY* is also a kind of hierarchy) is probably that in the Hierarchy the processing of information takes place at two different levels subsequent to each other in time, but in the Edge there is only one processing level. In the Hierarchy there is first some processing within the functional team (e.g., the *who* team). This processing involves the team members and their team leader. Next, there is a second sequence of processing, involving the team leaders for all four teams and the cross team coordinator, (i.e., the commander). This takes time and moreover, we saw examples how the team-leaders did not share the threat evaluation made by their team members and thus did not report the team evaluation to the cross team coordinator but only their individual assessment, resulting in the command team getting it wrong while the functional team had the correct assessment.

If we compare the results with our expectations the only expectation that was not supported by the results was that the Edge should be slower than the other two types to make its consensus decision. We assumed that it would take longer time to discuss (CHAT) among the members of the Edge in order to line up everyone to agree on a similar interpretation of the threat situation. This however showed to be not correct, because the Edge teams needed less internal discussions and time (through CHAT) than the other two organization types before they could agree on a decision.

An interesting finding that was unexpected was the emergent leadership roles that certain members took up as the Edge runs progressed (as opposed to the outlier Edge team who appointed a team-leader before the experiment run started). This may highlight the need for

allowing certain leadership and hierarchical structure (however flat this may be) to emerge in order for Edge organizations to collaborate and perform decision making via consensus.

Through detailed analysis of the SHARE and POST traffic, we noticed that regardless of the organizational concept, there were always players who had a higher propensity to share information and vice versa, and this may just reflect the variations in the human nature attitude in this respect. However, it may be relevant to ask questions about the effect of organizational culture and behaviour on organizational performance, which may or may not be an overriding factor as compared to the organizational structure itself. Or perhaps recognizing that it may be highly unlikely to achieve perfect organizational culture and behavior, it would be interesting to identify structures and concepts that are most robust against such variability.

Through detailed analysis of the overall CHAT traffic volume, we also noticed that the overall volume of CHAT traffic in the Edge runs are much lower than those in the Hybrid and Traditional Hierarchy runs. This is somewhat counter-intuitive as we expected that traffic volume to be higher in the Edge runs because there would be more discussion, debate and voting on the final results. The CHAT traffic in the *TH* and *HY* runs showed that there were duplicated CHAT entries between the functional chat rooms and the leader chat rooms as the team leaders had to transfer the discussion by the team members to the CHAT room and clarify with the other team leaders. This point to the finding that segregating the participants into separate CHAT rooms would result in much higher CHAT traffic but it need not translate to better results. In fact, the final results showed that congregating into one single CHATROOM, as is the set-up under the Edge runs, was the most efficient way of communicating within the group.

Conclusion

In our concluding segment, we would like to discuss issues pertaining to reliability and validity of our results. Firstly, as the number of observations is small, at only two data points per condition, this makes it difficult to establish how stable these results will be if the study were to be replicated. Secondly, as inferred from the fact that in one of the *HY* teams as well as in the outlier *E* team, a key factor for success in ELICIT, regardless of organizational structure, is that all (critical) information received by a player gets posted on a website so that more people can take account of it. This depends on the individual player's behaviour during the experiment and may not be attributed to any particular C2 structure. Having insufficient runs to average out this uncontrollable variance may have distorted the total results. However, this risk is affecting only the conclusions regarding the effectiveness of the Hybrid because one of those teams did not manage to disseminate all critical factoids (as also did not the outlier Edge team but that team was not included in the analyses). Thus, the effectiveness of the *HY* organization might be underestimated based on our results.

So what can we actually learn from this study in terms of C2 type organization efficiency? To address this question, the authors would like to discuss whether the results from this study can be generalized to real C2 organizations. We showed that in at least some circumstances consensus decision making in a flat organization does not have to take longer (time) than Hierarchical decision making, at least not in a situation where a true, or objectively correct, decision can be found. On the contrary, it can be faster. We also showed that filtering of information through Hierarchical "filters" is risky and sensitive to distortion in some cases. Some mid-level managers can suppress the opinion of their subordinate team members and present only their own personal view, and there is little or no incentive for a superior commander to surpass his subordinate commander and go directly to the team in order to get their view, especially if there was no common communication channel or medium available (such as a com-

mon website or chat room). We also showed that understanding of the intent, however simple it may be, (in this case to answer four different questions), was not enough. It is equally important for success that the individual entities of an organization (a) understands its own role (here to disseminate incoming information) and (b) that there is a functional working procedure in the organization so that all team members can contribute effectively (cf., Jensen 2006). Although not quite statistically robust, these findings do possess a fair amount of face validity.

Another issue is of course whether an Edge C2 organization would generally outperform a Hierarchical C2 organization regarding decision speed, decision accuracy, and level of shared (correct) situation awareness. This can not be concluded from our results, partly because of the reliability issues related previously, but mainly because the differences between the ELICIT and a real C2 organization are substantial. First, an organization normally has to take some action and not only make a decision, as in ELICIT. In real C2 situations, the need for action often requires some prioritizing of resources as well as determining the order in which the actions are taken. Such actions could prove fairly difficult and time consuming in an Edge organization. Second, in real C2 situations no objectively correct "truth" can ever be found, as in ELICIT. It is always a matter of opinion or opinions among the members of the organization on how a specific task and situation should be interpreted and real values are at stake, which make people more prone to fight for their beliefs. This is also a complicating factor for a flat organization. Third, in real C2 situations the participants tend to have a different level of experience and background, which makes them less able to act as equals in a flat organization.

However, ELICIT offers opportunity to study the effectiveness of different C2 structures in several areas, even more easily in the latest version as some of the restrictions in the set-up of the game now have been removed (cf., Ruddy and Nissen 2008). In future studies it would be interesting to study how representatives of different orga-

nizations come to consensus, and it would also be interesting to analyze further how leadership within the edge organization emerges spontaneously, and which individuals will step forward and take the lead in a team. In this current study reported here, we also measured individual decision-making styles of the participants and we plan to evaluate this further in order to find out if the emergent leaders of the Edge teams also have a common decision making style.

References

Alberts, David S., John J. Garstka, and Fredrik P. Stein. 1999. *Network Centric Warfare*. Washington, DC: CCRP Publication Series.

Alberts, David S. and Richard E. Hayes. 2003. *Power to the Edge: Command and Control in the Information Age.* Washington, DC: CCRP Publication Series.

Alberts, David S., and Richard E. Hayes, 2006(a). *Planning: Complex Endeavors*. Community Review Version (www.dodccrp.org). Washington, DC: CCRP Publication Series.

Alberts, David S., and Richard E. Hayes. 2006(b). *Understanding Command and Control*. Washington, DC: CCRP Publication Series.

Berube, Christopher D., Janet M. Hitzeman, Roderick J. Holland, Robert L. Anapol, and Stephen R. Moore. "Adapting C2 to the 21st Century." *Proceedings to 12th International Command and Control Research and Technology Symposium*, Newport, RI, June 19-21, 2007.

Brehmer, Berndt. 1998. "Effects of time-pressure in fully connected and hierarchical architectures of distributed decision-making." In Y. Wærn (Ed). *Co-operative Process Management: Cognition and Information Technology*. London; Bristol, PA: Taylor and Francis.

Brehmer, Berndt. "Command without Commanders." *Proceedings to 14th International Command and Control Research and Technology Symposium*, Washington, DC, June 15-17, 2009.

Builder, Carl H., Steven C. Bankes, and Richard Nordin. 1999. "Command concepts: A theory derived from the practice of command and control." Santa Monica, CA: RAND.

Cheah, Mervyn, C. Ngoh, Gwenda Fong, and Elsie Toh. "NCW in Action - Experimentation within A Distributed and Integrated Command Environment (DICE)." *Proceedings to 12th International Command and Control Research and Technology Symposium*, Newport, RI, June 19-21. 2007.

Cohen, John. 1977. *Statistical Power Analysis for the Behavioural Sciences*. Academic Press: New York.

Cohen, John. 1992. A Power Primer. *Psychological Bulletin*,112: 155-159.

Dekker, Tony. "Centralization vs. Self-synchronization: An Agent-Based Investigation." *Proceedings to 11th International Command and Control Research and Technology Symposium,* Cambridge, UK, September 26-28, 2006.

Duncan, Matthew, and Marie-Eve Jobidon. "Spontaneous Role Adoption and Self-synchronization in Edge Organizations Using the ELICIT Platform." *Proceedings to 13th International Command and Control Research and Technology Symposium*, Seattle, WA, June 17-19, 2008.

Dynes, R. R. 1994. Community emergency planning: False assumptions and inappropriate analogies. *International Journal of Mass Emergencies and Disasters*, 12: 141-158.

Eovito, Bryan A. "The impact of Synchronous Text-Based Chat on Military Command and Control." *Proceedings to 2006 Command and Control Research and Technology Symposium*, San Diego, CA, June 20-22, 2006.

Hanlon, Edvard. 2004. "Distributed Operations: The Time is Now." Marine Corps Gazette, July, pp. 36-54.

Jensen, Eva. "Good sensemaking is more important than information for the quality of plans." *Proceedings to 11th International Command and Control Research and Technology Symposium*, Cambridge, UK, Sept. 26-28, 2006.

Leweling, Tara A., and Mark. E. Nissen. "Hypothesis Testing of Edge Organizations: Laboratory Experimentation Using the ELICIT Multiplayer Intelligence Game." *Proceedings to 12th International Command and Control Research and Technology Symposium*, Newport, RI, June 19-21, 2007.

Martin, Danielle M., and Jimmie C. McEver, III. "Metrics, Analysis and Methods for Exploitation of ELICIT Experimental Data." *Proceedings to 13th International Command and Control Research and Technology Symposium*, Seattle, WA, June 17-19, 2008.

NATO SAS-065 Research Task Group, "NATO NEC C2 Maturity Model Overview" (working publication 7 August 2009, Washington, DC: CCRP Publication, 2009).

Parity Communications Inc. 2006. Experiments in Command and Control within Edge Organizations. Final Report for subcontract EBR06-0002 Evidence Based Research. Washington, DC: Command and Control Research Program (CCRP).

Rigas, Georgios, Mats Persson, and Berndt Brehmer. 2005. N*etwork Centric Warfare: Final Project Report* (in Swedish). Stockholm, SE: Swedish National Defense College.

Ruddy, Mary. "ELICIT – The Experimental Laboratory for Investigating Collaboration, Information-sharing and Trust." *Proceedings to 12th International Command and Control Research and Technology Symposium*, Newport, RI, June 19-21, 2007.

Ruddy, Mary, and Mark Nissen. "New software Platform Capabilities and Experimentation Campaign for ELICIT." *Proceedings to 13th International Command and Control Research and Technology Symposium*, Seattle, WA, June 17-19, 2008.

Schmidle, R. E., and F. G. Hoffman. 2004. "Commanding the Contested Zones." Marine Corps Gazette, July.

Simpkin, Richard E. 1985. *The Race to the Swift: Thoughts on twenty-first century warfare*. London: Brassey's Defense Publishers.

Simpson, Marvin L. "User's Epistle on Text Chat Tool Acquisition." *Proceedings to 2006 Command and Control Research and Technology Symposium*, San Diego, CA, June 20-22, 2006.

US Army. 2003. *Field Manual 6-0 Mission Command: Command and Control of Army Forces*. Headquarters, Department of the Army: Washington, DC.

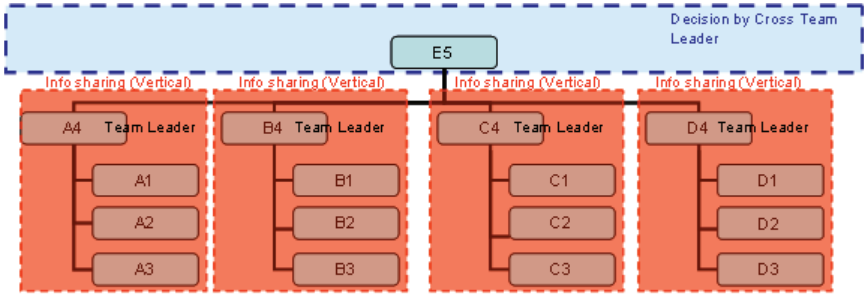
US Army. 2005. *Field Manual 5-0 Army Planning and Orders Production*. Headquarters, Department of the Army: Washington, DC.

Wilbeck, Christopher W., 2003. *Command of the Objective Force*. United States Army Command and General Staff College, Fort Leavenworth, KS. School of Advanced Military Studies. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA416216&Location=U2&doc=GetTRDoc.pdf

APPENDIX I

EXPERIMENTATION MATERIALS

Description of the Traditional Hierarchy (TH) Concept



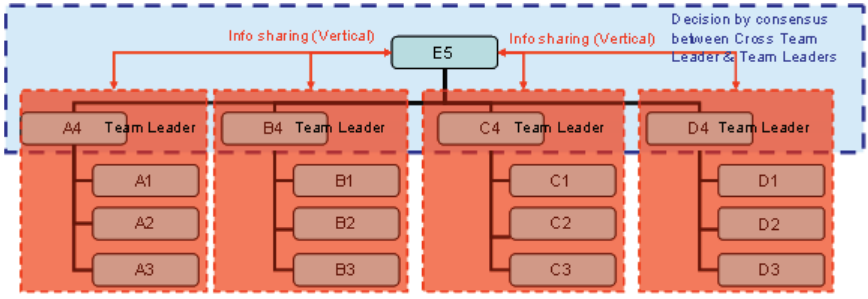
C2 Features:

No. of Layers	**Three: Cross Team Leader, Team Leader, Team Member**
Grouping	Organised in Functional Groups: Who, What, Where, When
Information Access	Access to websites of OWN functional group
Communication Channels through CHAT	Members can CHAT with Members. Members can CHAT with their own Team Leader. Team Leaders can CHAT with other Team Leaders and Cross Team Leader. Members cannot CHAT with Cross Team Leader.
Decision Making	Only by Cross Team Leader

Briefing Instructions:

Item	Instructions
Grouping	▪ There will be 4 functional groups. ▪ Each person will be assigned to a functional group as either leader or members. ▪ There is a cross-team leader on top of the 4 functional groups.
Information Access	▪ Each member and team leader have access his OWN functional websites only ▪ Cross-team Leader has access to all functional websites.
Communication Channels through CHAT	▪ Each member can communicate with his OWN team member and OWN team leader through the respective WHO/WHAT/WHERE/WHEN chat room. ▪ Each member can also communicate with members of other functional groups through MEMBER chat room. ▪ Each team leader can communicate with team leaders of other functional groups and the cross-team leader through TEAM LEADER-CROSS TEAM LEADER chat room.
Decision Making	▪ Final answer to be decided by Cross-Team Leader. ▪ When Cross-Team Leader has submitted the final answer, Controller will ask everyone to submit whatever answers they feel are correct.

Description of the Hybrid (HY) Concept



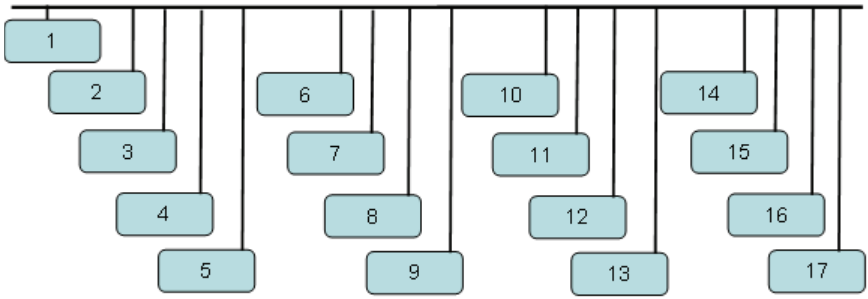
C2 Features:

No. of Layers	**Three: Cross Team Leader, Team Leader, Team Member**
Grouping	Organised in Functional Groups: Who, What, Where, When
Information Access	Access to ALL FOUR functional websites
Communication Channels through CHAT	Everyone can CHAT with everyone else.
Decision Making	Consensus by Cross Team Leader with Team Leaders

Briefing Instructions:

Item	**Instructions**
Grouping	▪ There will be 4 functional groups. ▪ Each person will be assigned to a functional group as either leader or members. ▪ There is a cross-team leader on top of the 4 functional groups.
Information Access	▪ Each member and team leader have access ALL functional websites. ▪ Cross-team Leader has access to ALL functional websites.
Communication Channels through CHAT	▪ Each member can communicate with his OWN team member and OWN team leader through the respective WHO/WHAT/WHERE/WHEN chat room. ▪ Each member can also communicate with members of other functional groups PLUS the cross-team leader through MEMBER chat room. ▪ Each team leader can communicate with team leaders of other functional groups and the cross-team leader through TEAM LEADER-CROSS TEAM LEADER chat room.
Decision Making	▪ Final answer to be decided by consensus among leaders, ie. 3 leaders out of 5 leaders/cross-team leader. ▪ If any leader thinks he has the correct answer, he should seek 2 more consents from the other leaders. ▪ If he obtains a total at least 3 consents, including himself, he should inform the other leaders that he has sufficient majority and seek support to submit the final decision. ▪ If he obtains at least 3 supports, he should submit the answer. ▪ When the final answer is submitted, Controller will ask everyone to submit whatever answers they feel are correct.

Description of the Edge (E) Concept



C2 Features:

No. of Layers	One layer for all
Grouping	Nil
Information Access	Access to ALL FOUR functional websites
Communication Channels through CHAT	Everyone can CHAT with everyone else.
Decision Making	Consensus by all members

Briefing Instructions:

Item	Instructions
Grouping	▪ There will be NO functional groups and leaders. ▪ Everyone is a member. ▪ Each member is free to choose to work on any one or more areas.
Information Access	▪ Each member have access ALL functional websites.
Communication Channels through CHAT	▪ Each member can communicate with any member through the COMMON chat room.
Decision Making	▪ Final answer to be decided by MAJORITY among members, i.e., 9 out of 17 members. ▪ If any member thinks he has the correct answer, he should seek 8 more consents from the 16 other members. ▪ If he obtains a total at least 9 consents including himself, he should inform the rest that he has sufficient majority and seek support to submit the final decision. ▪ If he obtains at least 9 supports, he should submit the answer. ▪ When the final answer is submitted, Controller will ask everyone to submit whatever answers they feel are correct.

Description of Experiment Procedure

Programme for conduct of experiment:

Item	Description
Introduction Brief	Participants were given a description of the 3 experimented C2 concepts, i.e., Traditional Hierarchy, Hybrid, and Edge.
Introduction to ELICIT and CHAT	Participants were given a hands-on training of the ELICIT game and the CHAT.
Conduct of Training Run	Participants went through a complete run using one of the 3 experimented C2 concepts. Each team was given a C2 concept that would be different from the concept that the team would be adopting in the Actual Run. The intention was to allow participants the opportunity to familiarise with the ELICIT game and CHAT functionalities through an actual complete run without getting too proficient with the C2 concept. By doing this, we hoped to removed the time taken for the participants to familiarise on the ELICIT game and CHAT functionalities in the Actual run.
Briefing of Actual Run	The C2 concept that each team was suppose to adopt in the Actual Run was revealed. The teams were also reminded of the details of the structure and C2 features, complete with the instructions.
Conduct of Actual Run	All teams completed their run using the same scenario in ELICIT, but the scenario was a different one from the training run. Controllers were positioned to observe the development of discussions in CHAT and the behaviour of the participants during the runs. Each run was terminated when the final answer was submitted and each player had submitted their own understanding of the answer.

Assignment of C2 concepts to teams:

Team	1	2	3	4	5	6	7
Training Run	TH	HY	TH	E	TH	E	HY
Actual Run	E	E	HY	HY	E	TH	TH