

ELICIT – The Experimental Laboratory for
Investigating Collaboration, Information-sharing
and Trust

April 2, 2007

Prepared By
Mary Ruddy
Parity Communications, Inc.

Table of Contents

ELICIT Overview	3
Background	3
Meeting the Core Requirements	4
Approach.....	4
Experiment Software Platform	5
Experiment Design	5
Purpose	5
Approach	6
Details of Experiment Design	9
Experiment Variables and Metrics.....	14
Experiment Protocols.....	15
Data Collection Process and Analysis Approach	15
Preliminary Results and Conclusions	15
Appendixes.....	22
A: Summary of Experiment Model	22
B: Experiment Variables and Metrics	26
C: Instructions for Edge (A) group	30
D: Instructions for C2 (B) group.....	32
E: Experiment Procedure	35
F: Experiment Software.....	36
G: Installing the ELICIT Software.....	48
H: Factoid Sets and Factoid Keys	50
I: ELICIT Action Log.....	55
J: Creating Data Analysis Files — Post-Experiment Analysis	60
K: Running Experiments in a University Setting.....	66
L: Suggestions for Future Work.....	70
M: Web Site Support	74
N: Resources	75

ELICIT Overview

Background

In Power to the Edge, Alberts and Hayes (2003) argue that missions designed with superior shared awareness, trust and self-synchronization will perform with greater speed, precision, effectiveness, and agility than missions conducted under traditional hierarchical command structures. They further argue that this is achieved by placing decision rights at the “edge of the organization,” close to the points of consequence.

As part of its network-centric warfare initiative, the Command and Control Research Program (CCRP) is engaged in developing and testing principles of organization that significantly revise traditional command and control practices, transferring power and decision rights to the edge of the organization. In order to test these assertions, CCRP needs to frame testable hypotheses about the relative effectiveness of edge organizations in comparison to other methods of organization through a series of real-world experiments.

In order for CCRP to undertake such experiments, the following capabilities are needed:

Replicable and valid measures of shared awareness, self-synchrony and trust. Ideally, such metrics need to be derived from observed behaviors in organizational settings.

Non-intrusive instrumentation that can be used to capture real-time behavioral metrics about different types of organizational interactions.

Automated tools and techniques that can accommodate testing for the different factors that might affect “edge performance” with respect to shared awareness, trust, and self-synchrony.

Privacy controls and methods of apparatus deployment and administration that do not themselves entail significant organizational changes or overhead.

Another major requirement for conducting such experiments is that they compare the relative effectiveness of edge organization to traditional command and control

ICCRTS - ELICIT

principles in missions that either mimic real world missions or are themselves actual missions.

By creating and evolving the ELICIT software platform, tools and procedures, and by conducting ELICIT experiments we seek to test the validity of these propositions in a controlled environment and within a controlled task domain. The objective of the experiment design is to conduct a series of online experiments to compare the relative efficiency and effectiveness of traditional *command and control (C2)* vs. self-organizing, peer-based *edge (E)* organizational forms in performing tasks that require decision making and collaboration.

Meeting the Core Requirements

The major efforts in this project included the design, development and testing of the software platform on which to run the experiments; the design and execution of the experiment task (including supporting materials); and data analysis.

This project addressed the following specific objectives:

Develop system-based behavioral measures of “shared awareness.” By controlling for the distribution of content and its visibility, and capturing in time logs when different subjects have shared awareness, the project was designed to compare shared awareness among subjects and its impact upon the successful completion of a mission.

Develop system-based behavioral measures of “trust” by monitoring subject interactions in terms of reciprocity, responsiveness, number of interactions, and willingness to share content. Such trust measures can be used as predictors of mission effectiveness and timeliness.

Develop subject-based indicators of “self-synchronization” based upon the effectiveness of trust and shared awareness in reducing decision cycles.

Approach

The ELICIT software was developed and iteratively refined using live subjects. The experiments are controlled-hypothesis testing experiments. The experiment task is to identify the who, what, where and when of an adversary attack based on simple information facts (called “factoids”) that become known to a team. The independent variable is whether a team is organized using traditional command and control hierarchy or using edge organization principles.

Experiment Software Platform

The ELICIT software is a downloadable software application that is installed on each subject's laptop. It was built on top of the open source Higgins Trust Framework software developed as part of the SocialPhysics project affiliated with the Berkman Center for Internet and Society at Harvard Law School. The software allows CCRP and other experimenters to precisely model specific C2 processes, as well as edge organization processes, and to fully instrument all interactions.

The ELICIT software platform includes a measurement capability built over a messaging infrastructure. Unlike existing software messaging and analysis technology, this open source software was uniquely designed to enable shared awareness, trust building and self-synchronization. The ELICIT platform was designed to be configurable to support both initial and follow-on experiments.

The software, which is built on the Eclipse Rich Client Platform (RCP), offers modular, plug-in based design. Thus it is relatively easy to modify the software to support different experiment features or to add further communications mechanisms to determine their impact on team efficiency. Details on the experiment software and its capabilities are included in Appendix F.

Experiment Design

Purpose

The objective is to conduct a series of online experiments to compare the relative efficiency and effectiveness of *command and control (C2)* organizational structure with a networked, peer-based *edge (E)* organization in performing tasks that require decision making and collaboration.

One of the key propositions of edge organization is that unknown parties, when given a shared awareness, will collaborate and self-synchronize their behaviors to achieve common goals. It is also argued that by placing decision rights at the "edge of the organization," close to the points of conflict and consequence, more efficient and effective decisions can be made than when decision rights are concentrated and

ICCRTS - ELICIT

controlled through a hierarchical command and control structure. (Alberts and Hayes, 2003)

The experiment challenge is to provide an online test bed where the efficiency and efficacy of command and control and edge organizational models can be compared for tasks that mimic real world conditions and challenges. By isolating several of the key structural and interactive factors that characterize command and control and edge organizations, a series of experiments and measures can be constructed to compare their relative effectiveness and efficiencies for an identical set of tasks.

The controlled hypothesis testing experiments were designed to further understanding of the advantages of edge organization that are being conducted as part of much larger efforts in investigating the applicability of edge organization.

Approach

The approach taken was to design an online multi-subject task that is sufficiently rich that it mimics issues in real-world conditions. At the same time it needed to be simple enough to make it possible to control for multiple structural and interactive parameters that differentiate between command and control-based and edge-based organizational models.

The experiment challenge was to construct an experiment task that had the following characteristics:

- Highly relevant to real-world situations
- Of current interest to the DoD community
- Of interest and engaging to experiment subjects
- Multi-user, with each subject treated equivalently
- Abstract enough that it can be fully modeled
- Capable of being constructed in several versions (to support practice round and multiple rounds of experiments with the same subject group.)
- Short enough so that both the practice round and actual round can be conducted in a reasonable amount of time
- Difficult enough that it is not trivial to accomplish regardless of the organization structure used
- Comprehensive in its range of difficulties, so that the effects of organizational structure can be seen

ICCRTS - ELICIT

- Flexible, so it can be adapted to new settings by changing names, tables, etc.

By developing the ELICIT software platform, we were able to precisely model specific C2 and edge organization processes, and instrument all interactions. By abstracting complex tasks to simple interactions we were able to control for variations in subject capability and skill, and for the inherent variability in live situations that have confounded the data in other experiments. This approach allows the experiments to focus on fundamental issues of organizational design and to isolate factors that can be used to improve efficiency and effectiveness. The modular software can be modified to test additional approaches in future experiments.

In the experiment, subjects are randomly assigned to two groups: “Edge” (*E*) and “Command and Control” (*C2*). Each group is to identify the who, what, where and when of an adversary attack by combining and sharing a set of information factoids that are distributed among the subjects.

There are four kinds of factoids corresponding to the four kinds of information required (who, what, where and when). Like pieces of a puzzle, each contains a piece of information, but each alone is insufficient. In the experiments conducted to date, all factoids were factually correct; no incorrect information was used.

Since subjects have only partial information, they must collaborate and exchange information with other subjects in their group in order to complete the task. All interactions between subjects occur through a software application resident on each subject’s computer. In either group, any subject can communicate with any other subject, although all communication occurs *only* as mediated by the software application.

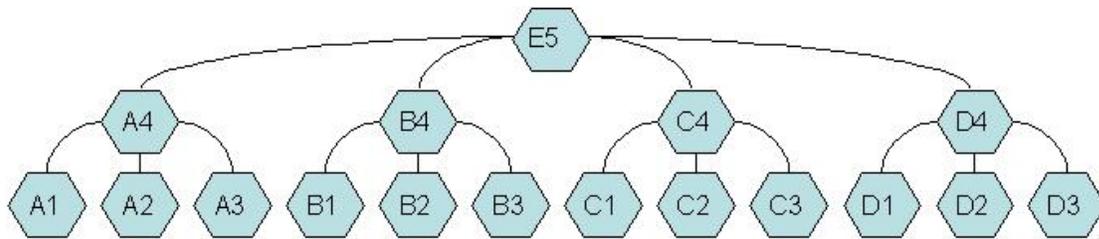
The experiment software monitors the progress of the information gathering task and declares the trial over when each individual has identified the who, what, where and when – or when the experiment times out. The group of subjects (*E* or *C2*) that completes the task first is declared the “winner.”

In order to minimize any side effects from variations in previous knowledge of subject matter or the ability to absorb subject information, the information in the task is highly abstracted.

The experiment is fully instrumented by the ELICIT software. It records the time and particulars of every action by each subject.

Organization of the C2 Community

In the C2 organization, there are four teams of four members each plus an overall cross-team coordinator (designated in the chart as E5). The four teams are organized along the lines of a traditional hierarchical command structure, each with a leader. The following shows the hierarchical relationship between the overall coordinator (E5), the four leaders (A4, B4, C4 and D4) and their subordinates:



Team A is from country A. Teams B, C and D are from countries B, C and D respectively. The four teams (A, B, C and D) each have a functional specialization: Team A is focused on who, team B on what, team C on where, and team D on when. The overall coordinator coordinates information among the team leaders across team boundaries.

NOTE: The above diagram is an *organization* chart; it is not a description of the possible communication flows. For example, subjects are able to share information to any other subject, not just along the lines shown above.

Organization of the E Community

In the E community the subjects are organized along edge principles. Unlike the C2 organization, there is no hierarchical decomposition nor is there specialization by functional area. Decision rights are decentralized: subjects decide for themselves what aspect of the task to work on, and in some situations they can choose other subjects to work with. Control is achieved entirely through the shared awareness provided by universal access to information on shared information systems.

Additional details about the organizational factors modeled for the experiment can be found in Appendix A.

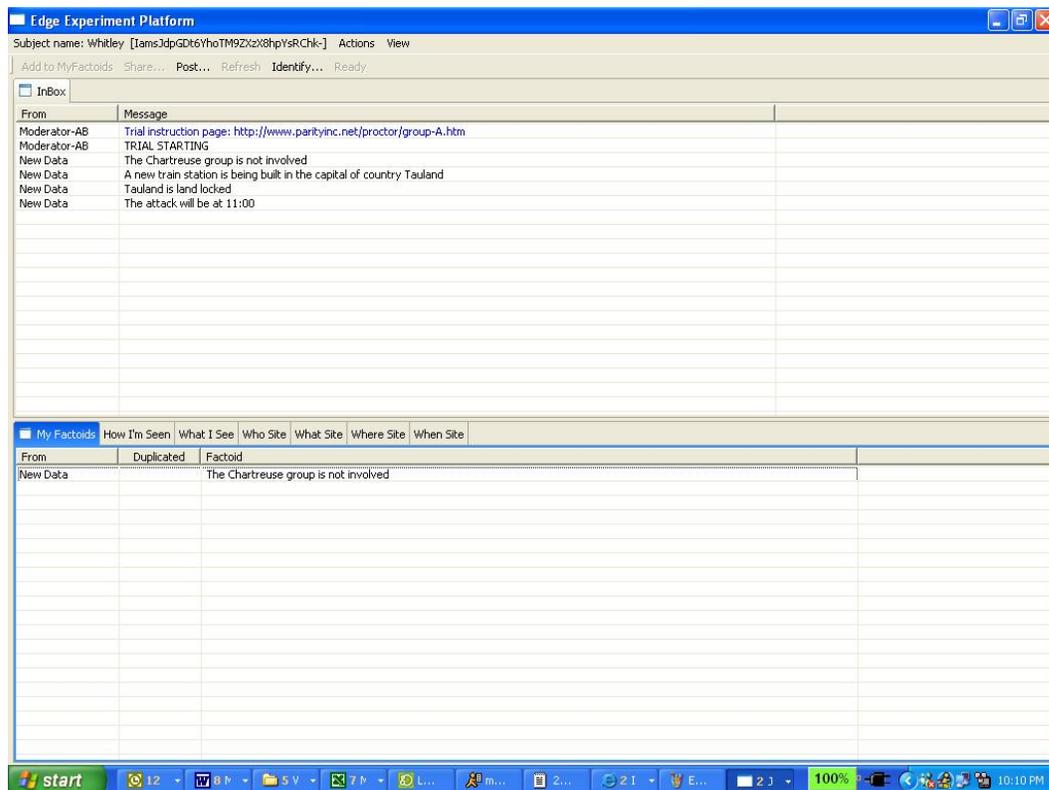
Details of Experiment Design

This section provides additional details of the experiment design.

Subjects

Subjects are randomly assigned to two 17-member groups, C2 and E. The subjects are asked to perform the task using a software application that has been pre-loaded onto their computers.

All experiment communications occur between anonymous identities, so previous relationships between subjects are irrelevant. The subjects need not be physically in the same location, as all of their interactions are mediated through the software application.



Subjects see and use a simple screen that contains:

- A message queue (looks like an email inbox) that displays messages from the moderator as well as factoids (which look like one-line email messages)
- A multi-tabbed information display area that displays information about other organizational members as well as simplified Web site-like lists to which subjects can post factoids they have received

ICCRTS - ELICIT

- A set of actions (menu items) the subject can take. The important actions are: (i) *Sharing* a factoid with another subject; (ii) *Posting* a factoid to one of the Web site-like lists; (iii) *Pulling* to see what is on a Web site-like list; and (iv) *Identifying* one or more aspects of the adversary attack (when the subject thinks they know some or all of this information).

By tightly constraining the forms of communication, the experiment controls for variations in communications styles among the subjects.

Task Objective

The subjects are given the objective to solve the puzzle of the location, time, target and group responsible for the adversary attack. Instructions are tailored to fit either the E organization (designated the “A” group) and its roles or the C2 organization (designated the “B” group). All subjects are instructed that they are free to work on any aspect of the task.

Subject Roles

The experiment software automatically provides a URL with group-specific instructions to each subject at the start of an experiment trial set.

E Group-Specific Instructions and Capabilities

The latest version of the base Edge group instructions Web page is available at <http://www.dodccrp.org/html3/elicite.html> and is provided in Appendix C.

C2 Group-Specific Instructions and Capabilities

The latest version of the base C2 group instructions Web page is available at <http://www.dodccrp.org/html3/elicite.html> and is provided in Appendix D.

Factoids

During each round of the experiment, the application delivers four “factoids” to the inbox of each subject. Group members use these factoids to help identify the adversary attack. Two factoids arrive at the beginning of the experiment round. An additional factoid is distributed five minutes later and one last factoid five minutes after that. All of the information necessary to identify the adversary attack has been

ICCRTS - ELICIT

distributed to the subject group within the first 10 minutes after the start of an experiment round.

There are four types of factoids that represent information about the anticipated attack:

- Who factoids – the likely actors
- What factoids – the target
- Where factoids – the country
- When factoids – the month, day and time.

Initial Factoid Distribution

There are 68 factoids, including 4 expertise factoids.

To model a hierarchy (C2) organization, in which team leaders traditionally have more expertise than their team members, each of the four team leaders is given an “expertise” factoid that represents pre-attained knowledge. In the E case, these expertise factoids are disbursed within the community at random.

Some of the 68 factoids are more important than others. The more important ones are considered *key* factoids. Expertise factoids are special key factoids. The remaining non-key factoids are distributed among the subjects so as to ensure that no subject receives more than 1 key factoid. The base factoid sets are available at <http://www.dodccrp.org/html3/elicit.html>.

This phased factoid distribution is designed so that the task can't be solved until the last distribution is made.

Distribution of factoids is controlled and specified in factoid set tables. Four complete factoid sets were created. Factoid set 4 is the easiest. Factoids sets 2 and 3 are very similar. For the initial experiments, all live subject tests were performed using factoid set 4-17 for the practice round and 1-17 for the actual round.

In subsequent trials, factoid sets and or players can be changed. As the factoids are stored in tables, it is possible to create additional factoid sets. The Naval Post Graduate School in Monterey has already created and used derivative ELICIT factoid sets.

See Appendix H for additional information about factoids, factoid sets and factoid keys.

Experiment A – Civilian Educational Organization

In this experiment, administered at a private university, the following definition and null hypotheses were used.

Definition of Performance: Performance is measured by the average time and average number of actions the team's subjects require to complete the task.

Null Hypothesis 1: Measures of self-synchronization do not correlate with performance.

Null Hypothesis 2: There is no difference in performance in achieving a mission between those subjects who have increased shared awareness and those that do not.

Null Hypothesis 3: The platform-derived measures of trust and shared awareness do not correlate with the results of these same measures using traditional social network analysis of survey-gathered data.

Null Hypothesis 4: There is no difference in performance in achieving a mission between those groups with a command and control organization and those groups with an edge organization.

Experiment B – DoD

Our pre-test runs were done at an organization performing government research. Additional runs are in process at a military academy. This data will be used to begin to address Null Hypothesis 5.

Null Hypothesis 5: There is no difference in performance in achieving a mission between those subjects in DoD and those subjects in educational organizations.

Experiment Trials

Dependent Variables	Trial
---------------------	-------

Organization Structure		
Non-DoD	C2	1b
Non-DoD	E	1a
DoD	C2	2b
DoD	E	2a

Testing Hypothesis 1 (H1)

During these experiments, members of the test group were asked to participate in a series of simple questionnaires following Social Network Analysis (SNA) methodology. In this approach individual members respond to survey questions that are designed to surface a particular kind of links within the network. A series of questions was designed to elucidate these network linkages from a variety of perspectives. The survey questions were initially developed by Patti Anklam who is a practicing Social Network Analysis consultant. Her new book is listed in Appendix N. The survey had several sections with differing objectives:

- Provide general demographic data that could be used to explain variations in subject performance.
- Provide data to perform formal SNA analysis, as helpful and applicable.
- Provide additional information on shared awareness. (e.g. changes in understanding since the last Identify action was taken.
- Provide general feedback on the experiment process.

The analysis phase consists of correlating the survey data with the data available from ELICIT. In this correlation, we expect to disprove hypothesis H1 and show that one can derive metrics from subject communications that are valid predictors of key edge organization capabilities and characteristics.

Testing Hypotheses 2-5 (H2-H5)

To test H2-H4, subjects were divided into experiment groups that use edge principles such as metrics of shared awareness, trust and self-synchronization, and control groups that do not. To test H5, we compare the results between groups in a university setting and the government/DOD sector. All groups were instrumented with ELICIT and its metrics were used to make comparisons.

Experiment Variables and Metrics

Details on the experiment variables and metrics are shown in Appendix B.

Experiment Protocol

A full experiment trial consists of 34 subjects who are randomly assigned to either the Command and Control (C2) or Edge [E] team.

The experiment consists of four phases.

1. An introduction phase that includes:
 - An overview of the experiment's agenda, which is delivered via PowerPoint so that it can be adjusted for any logistical specifics.
 - An eight-minute subject pre-experiment briefing video, which is delivered via Windows Media Player.
2. A practice round (round 1), in which the subjects have the opportunity to use the software with a sample scenario. This round is designed to run for 20 minutes.
3. A second round, which uses a different scenario. This round is designed to run for up to an hour.
4. A wrap-up, which includes:
 - A Web-based survey, which requires 20 minutes to complete
 - A two-minute subject post-briefing video, which is delivered via Windows Media Player
 - Discussion of the experiment.

Total elapsed time for the four phases of the experiment is approximately 3 ½ hours.

Materials necessary to conduct the experiment – including pre-briefing videos, instructions for the experiment moderator, etc. – are available for download at the project's companion Web site <http://www.dodccrp.org/html3/elicit.html>. Additional

ICCRTS - ELICIT

information on experiment protocols is available in Appendix E. Suggestions for running *ad hoc* experiments in a university setting are discussed in Appendix K.

Data Collection Procedures and Analysis Approach

The principal means of data collection for the experiments are:

- Automated data collection that is integrated with the software used for the experiment
- A Web-based post-experiment survey that is automatically administered by the test software at the conclusion of an experiment trial
- Notes made by the subjects
- Observations/notes by the moderators and proctors.

The structure of the log files and a general approach to analyzing the raw log files are documented in Appendix I and Appendix J, respectively

Preliminary Results and Conclusions

To date ten runs have been made with live subjects, of which four are considered good runs. We are in the process of scheduling additional runs. Many more runs are needed before conclusive results can be derived. In addition, the Navel Post Graduate School at Monterey has made several dozen runs of a very interesting variation of this experiment based on the ELICIT platform. These results are being published separately. Ultimately this additional raw data will be made generally available under the ELICIT umbrella.

Preliminary results indicate that within a community, an Edge organization was more effective than a hierarchical organization at creating shared awareness and at task completion. As additional runs are made and the data analyzed, results will be posted to <http://www.dodccrp.org/html3/elicit.html>.

Task Difficulty

The goal had been to make a task that took a reasonable amount of time (e.g., one hour) and required inter-subject interactions. Additionally, the objective was to avoid making the purely cognitive aspects so difficult that the intelligence and information-managing aspects of the task would become the predominant criterion

ICCRTS - ELICIT

for success. Our intent was to measure organizational effectiveness, not individual IQ.

While it is possible to discuss task difficulty in the abstract (i.e., presuming complete knowledge of all factoids) in practice we found that even after a full hour (50 minutes after all factoids have been distributed), full diffusion of all factoids did not occur. For example, in the June 22 Edge trial, two important factoids were not posted. As a consequence, many subjects never received all the information necessary to correctly identify all aspects of the adversary attack.

Shared Awareness

One of the interesting results of the experiment is the degree to which information was not completely diffused. The diffusion varied based on the organization structure.

For example, in three of the experiment runs after the debriefing phase, one of the subjects began discussing the fact that another subject had hoarded a key piece of information. In one case there was a discussion about a subject hoarding a factoid to improve their score relative to the other subjects.

Given that each subject only receives 4 new factoids, and that posting those factoids to all available Web sites takes at most a total of 16 Post actions, it was interesting to observe that posting rates were so low (i.e., compared to the overall action activity rates). In the same period subjects queried these same Web sites on average at least once per minute looking for additional information.

For example, the <http://www.dodccrp.org/html3/elicit.html> Web site contains a *Percentage of Factoids Posted Workbook*. The data shows that for the runs analyzed, at the end of the first five minutes, only 71-79 percent of the factoids had been posted. By the end of an hour only 88 percent of the factoids had been posted in an Edge group, with only 69 and 74 percent posted in two Hierarchy group trials. For the runs analyzed, the percentage of factoids that was Posted was always higher for Edge vs. Hierarchy arrangements.

The *Activity and Reciprocation Workbook* shows that Edge groups had *pull* rates (accessing a Web site) that were approximately twice that of Hierarchy groups. This

ICCRTS - ELICIT

may be partly because they had four times the number Web sites that they could pull from.

Overall activity rates were high, showing that subjects were engaged. Pull rates were approximately one per minute for Hierarchy groups and two per minute for Edge groups over the entire one-hour period.

The use of the Share action was lower in Edge groups than in Hierarchy groups. For Edge groups Shares occurred only once a minute for the whole group. This makes sense, as Edge groups have more post options for communication.

Surprisingly in the Hierarchy scenario, where sharing is required for successful task performance, some of the subjects did not Share any of their factoids, even though others had Shared factoids with them.

Trust

Reciprocity did not seem to be an important factor. Preliminary analysis indicates that being the recipient of a Share (that is, receiving a factoid from another subject) was not a predictor of the recipient Sharing back. See the *Activity and Reciprocation Workbook*.

Preliminary analysis of the survey results for June 22 and June 23 showed that there was more trust developed in the Edge group than the Hierarchy group, where trust is measured as the ratio of the total number of subjects who received a “Would want in group” next time rating vs. a “Don’t want in group” rating.

- In the Edge round on June 22, there were 80 “Would want in group” ratings and 30 “Don’t want in group” ratings — a ratio of 2.67.
- In the Hierarchy round on June 23, there were 72 “Would want in group” ratings and 115 “Don’t want in group” ratings — a ratio of 0.63.

Note that other possible responses were neutral ratings such as “No preference.”

Anecdotally, after the conclusion of one of these rounds, two subjects were heard conversing about another subject who never Shared, and who they wouldn’t want to have in their group next time. Additionally, subjects who overshared or “spammed” other subjects with redundant information were also discussed negatively in post-experiment discussion.

Task Effectiveness

Since a subject can only correctly identify an adversary attack after they have access to factoids originally received by other subjects, the sharing of information — using either the Share or Post actions — is key to subject and team effectiveness.

Despite the fact that subjects are instructed that communicating factoids is necessary to performing the task, and the factoids a subject receives are clearly not sufficient to identify the adversary attack, factoids communication was below the optimal level for fully completing the experiment task.

Hypotheses

The original project Hypothesis were:

- H1. The platform-derived measures of trust and shared awareness do not correlate with the results of these same measures using traditional social network analysis of survey-gathered data.
- H2. There is no difference in performance in achieving a mission between those subjects who have increased shared awareness and those that do not.
- H3. There is no difference in performance in achieving a mission between those subjects who have increased trust and those that do not.
- H4. There is no difference in performance in achieving a mission between those subjects who self-synchronize and those that do not.
- H5. There is no difference in performance in achieving a mission between those subjects in DOD and those subjects in private corporations.

H1

Preliminary analysis of the survey data indicates that subjects did fill out the survey and indicate things like "Whitley wanted Harlan in her group" or "Taylor didn't want Harlan in their group." Analysis of platform-derived log files allows us to measure whether or not Harlan shared factoids with Taylor and Whitley. These did not always correlate. Nevertheless with the current level of analysis that has been done and the lack of sufficient data sets we cannot at present disprove this hypothesis. Our current estimate is that both more analysis of the processed experiment data on hand is needed, as well as additional trials.

H2

The pair of Edge and Hierarchy trials (performed on June 22 and June 23 respectively) at BU gave us data that preliminarily supports disproval of this hypothesis. In the June 22 trial 70 "units of shared awareness" produced 82 "units of performance", whereas in the June 23 trial only 14 "units of shared awareness" produced only 34 "units of performance." The Edge group had much greater dispersion of the factoids and was much more accurate in identifying the adversary attack.

H3

This null cannot be disproved. If we measure overall organizational performance by the aggregate number of identifies, we find that the Edge group was much higher performing (e.g., 82 vs. 34 units of performance). However, if we look at trust as it has been defined (related to reciprocity, responsiveness, number of interactions), we find that it is far lower in the Edge group. Subjects in the Edge group used the Share action far less than the Hierarchy. It appears the benefits of the greater shared awareness from the 17-person wide posting greatly reduced the need for one-to-one sharing.

H4

The definition of "self-synchronization" is an aggregate of the component measures of "trust" and "shared awareness." These are individually addressed in nulls H2 and H3 above.

H5

The Hierarchy group at a government analysis firm performed better than the university student C2 group; however, the individuals at the government firm were also older and more experienced. We look forward to comparing a military school to a civilian school. We are in the process of completing a set of runs at a military academy and hope to have results to compare to the civilian school in time for the conference. Existing data is not sufficient to disprove this hypothesis.

Other Observations

At the conclusion of the pre-tests with live subjects, subjects were very interested in discussing the experience, conditions under which their performance could be improved and how organization impacts performance. Several persons in the discussion felt that the experience of participating in the experiment and subsequent discussions would be useful as part of an education process on organizational design and dynamics as it brings abstract issues to life.

Discussions with Boston-area universities have also identified professors who are interested in this work and having it presented in one of their organizational design classes. We are running additional trials and further refining the experiment apparatus so that we are able to more precisely quantify performance differences between the E and C2 groups.

The ELICIT platform itself has proven to fill an important void in the arsenal of research tools. A number of US and NATO, educational and research organizations are in various stages of preparing and running experiments based on ELICIT.

The CCRP is making all ELICIT materials, raw data and results available on a public Web site at <http://www.dodccrp.org/html3/elicit.html>. To date, ELICIT experiment trials have been run with live subjects at a DC-based research facility, Boston University, the Naval Post Graduate School and West Point. NATO groups in Canada, Singapore and Portugal are planning ELICIT experiments.

The ELICIT program is still in its early stages. There are additional analyses to be performed and new experiments to be designed and run. We have been impressed and encouraged by the caliber of the persons who have begun working with ELICIT in the past year and the quality of the suggestions and planned being made for future work. Persons interested in using the ELICIT software platform or data should contact Dr. David Alberts at the CCRP. Persons interested in talking with a research organization that has extensive experience using the ELICIT platform and creating variations on the specific experiment documented here should contact Dr. Mark Nissen at the Navel Post Graduate School in Monterey, California.

Suggestions for Future Work appear in Appendix L.

Appendix A: Summary of Experiment Model

The following table summarizes the experiment model for the Command and Control (C2) community and the Edge (E) community.

C2 Factors	Experiment Model
Modeled	
<p>Hierarchical Decomposition</p> <p>Organizations are divided into separate units that report up to a central leadership.</p>	<p>The C2 organization is divided into four separate teams of four all coordinated by a sole central coordinator.</p>
<p>Specialization</p> <p>Organizational subgroups develop professional specialties.</p>	<p>Each of the four teams in the C2 organization specializes in a different type of information.</p>
<p>Command by Directive</p> <p>Direct commands control the organization.</p>	<p>Team leaders have special expertise that they can provide to their teams.</p>
<p>Information Hoarding</p> <p>Information stays near functional unit.</p>	<p>Each team's Web site is visible only to that team, not to the other teams.</p>
<p>Fixed Leadership (leadership by position)</p>	<p>Subject's roles (e.g., assignment to A4) are pre-determined at the start of the experiment and are permanent.</p>
<p>Decentralized Execution</p>	<p>The organizational design suggests that the sub-unit groups would focus execution on a particular area (e.g., "who").</p>
<p>More Centralized Decision Rights</p>	<p>The organizational structure suggests that A4, B4, C4, D4 and especially E5 roles have a centralized cross-team or cross-all-teams decision rights and execution tasks.</p>
Not Modeled	
<p>Optimization</p>	
<p>Prescribed Organizational Processes</p>	

Edge Factors	Experiment Model
Modeled	
<p>Networked Organization</p>	
<p>Organizational structure is completely flat.</p>	<p>17 subject peers</p>
<p>Information Sharing, Post/Pull</p>	
<p>As far as possible information should be posted first, then shared. Information flows predominantly horizontally and independent of chain of command</p>	<p>The team is visible. All four community of interest Web sites are shared and available to be "pulled" by all subjects. The subject must choose which of these Web sites to view. Subjects can also see their own information. Because these Web sites are decoupled from any particular organizational unit, posting to them can be thought of as horizontal and certainly independent of chain of command</p>
<p>Shared Awareness</p>	
<p>Edge organizations are supported by information systems that allow information to be tagged and posted and thereby made available to be pulled by any actor based on their context.</p>	<p>The information contained in each individual factoid is posted to one of four "community of interest" "Web sites" where it is available to any Edge subject.</p>
<p>Emergent Leadership</p>	
<p>Leadership by competence.</p>	<p>Reputation system option allows competence to be surfaced and exploited. (Not yet implemented.)</p>
<p>Emergent Control</p>	
<p>Initial conditions and information awareness allow control to emerge.</p>	<p>Instead of being told what to do, subjects are self-directed to work on most pressing areas. Also an overall shared sense of "how we're doing" is made possible by the four shared Web sites.</p>
<p>Enhanced Decision Rights at Edge</p>	
<p>Edge members are more empowered, less constrained to make decisions.</p>	<p>Edge subjects themselves can choose to work on the sub-problem they think needs the most help.</p>

Universal Factors	Experiment Model
<p>Modeled</p> <p>Individual Work</p> <p>Work activities unrelated to communication, coordination, reporting, etc. with others.</p> <p>Initial Information</p> <p>Information analysts start off with a small amount of initial clues as to the attack.</p> <p>Multiple Types of Information</p> <p>To connect the dots, multiple dimensions of information must be integrated.</p> <p>Free Choice of Whom to Interact With</p> <p>People choose with whom they will share information.</p> <p>Asynchrony</p> <p>People can simultaneously interact with any number of other individuals.</p> <p>Differing Response/Reaction Rates</p> <p>People work at different rates given the same cognitive tasks.</p>	<p>Subjects must look at their factoids, compare them with what is already known and communicate new information to others.</p> <p>Subjects start with only two factoids. These two are not sufficient to answer any of the four major questions about the attack.</p> <p>Four dimensions of information are involved (who, what, where, and when).</p> <p>Subjects can choose from a list of any other subject with which to initiate an interaction.</p> <p>Subjects can choose to provide information to other subjects in any order. Subjects can be involved in multiple sequential interactions.</p> <p>There is no forced synchronization of subject's actions. They can occur as slowly as the subject desires, or as quickly as subjects are willing and able.</p>
<p>Weakly Modeled</p> <p>(Lateral) Coordination Work</p> <p>Individuals must expend effort coordinating activities within overall information processes. These include branch, join, and flow coordination mechanisms to coordinate parallel and sequential work.</p> <p>Not Modeled</p> <p>Variable information trustworthiness</p> <p>Information may be uncertain or</p>	<p>Subjects have the ability to explicitly share information laterally (with other subjects individually) as well as <i>post</i> information to a larger group of subjects.</p>

ICCRTS - ELICIT

completely incorrect. Trustworthiness
may vary by the source.

(Lateral) information/knowledge transfer work

Tacit and explicit knowledge transfer
between people.

Appendix B: Experiment Variables and Metrics

The following table lists the factors that are tested for correlation (C) in the experiment, and those that are tested for no difference (D).

	C2	E	DoD	Non-DoD	DoD C2	DoD E	Non-DoD C2	Non-DoD E	Trust-platform	Shared awareness - platform	Trust-survey	Shared awareness-survey	Self-synchronization
Independent	C2	D											
Independent	E												
Independent	DoD		D										
Independent	Non-DoD												
Independent	DoD C2				D	D							
Independent	DoD E							D					
Independent	Non-DoD												
Independent	C2							D					
Independent	Non-DoD E												
Intervening	Trust-platform									C	C	C	C
Intervening	Shared awareness - platform												
Intervening	Trust-survey									C	C	C	C
Intervening	Shared awareness-survey										C	C	C
Intervening	Self-synchronization												C

Legend

C Test for correlation

ICCRTS - ELICIT

	Test for no
D	difference Time to complete task. Number of moves.

Dependent Variables

The dependent variable is *task effectiveness*, which is measured by:

Group time to complete the task – time elapsed between start and end messages in seconds.

The number of actions used for each group with breakout by action type. Rate is also calculated.

- Pull rate – extraction and count of all the pull records divided by trial time
- Share rate – extraction and count of all the Share records divided by trial time
- Post rate – extraction and count of all the Post records divided by trial time
- Identify – extraction and count of all the identify records

The number of correct answers in each category – Manual inspection of all Identify free text messages. (Using a multiple choice rather than a freeform text format was thought to bias results as it would require listing all possible actions that could provide information the subject did not already know. We should consider revisiting this decision as part of future efforts). Counts of partial and incorrect information will also be made.

Measure of richness of information known by the players (measured by survey). Manual inspection of all free text answers to awareness questions. Responses are evaluated using a multi-part scale:

- Blank
- Incorrect answer
- Partially correct answer
- Full correct answer
- Previously correct answer
- Correct answer with additional information that is partially correct
- Correct answer with additional information that is correct.

Intervening Dependent Variables

The intervening dependent variables are trust, self-synchronization and shared awareness, as described in the following subsections.

Trust

Considered in this context, trust is the willingness to work together, to take risks and share information. It is measured by:

Survey measure of:

- Who would you want/not want to have on your team again?
- Who would you share with next time?

Percentage of times that information sharing is reciprocated.

- Percentage of Shares where the *to* and *from* subjects in one Share are the same as the *from* and *to* subjects in a subsequent Share.

NOTE: The above two items are examples of areas that would benefit from further automation as part of a future effort.

Self-Synchronization

Self-synchronization in this experiment is measured by analyzing the type of information shared in each group, compared to the types of information not yet resolved. The hope/expectation was that, in the interest of the entire group winning, the E subjects will communicate on whatever “type” topic needs work the most (or at least still needs work) and thus the E community self-synchronizes.

The C2 organization also has multiple communications channels that subjects could use for self-synchronization after they’d completed their silo/teamed portion of the task.

- Number of Shares by group.
- Percentage of distributed factoids that have been Posted

Shared Awareness

The C2 teams each have “local” shared awareness via their local Web site, but are not directly aware of what’s going on with respect to other types of factoids. The E community, on the other hand has total shared awareness of what’s going on in each of the four types of factoids through access to all Web sites.

In this experiment shared awareness is measured by:

Percentage of total distributed factoids that have been posted at the beginning of phases 2 and 3 and at the end of the trial, by Web site.

Independent Variables

The initial ELICIT experiment is controlled by the following independent variables:

- Team's organizational structure (C2 vs. E)
- Subject's organization (DoD vs. non-DoD).

Uncontrolled Variables

The initial ELICIT experiment contains the following uncontrolled variables:

- Work rate – determined by measuring individual's work rate during the experiment (average action rate by client).
- Game playing experience – determined by survey of subjects
- IQ (To date, all participants have been college graduates or undergraduates.)
- Memory skills
- Difficulty of problem.

Appendix C: Instructions for Edge (A) group

"A" Group

You have been randomly assigned to the "A" group. During the trial you will be analyzing information factoids that are sent to you. You will also be able to **Share** information factoids with any participant of your group.

The experiment software application supports two ways of informing participants in your group about factoids you have received. You can **Share** a factoid directly with another participant using the share menu item or you can **Post** a factoid to one of four Web sites. There are separate Who, What, Where and When sites. Participants can post intelligence information [factoids] they have to a Web site that others in their group can simultaneously see and update. Though these areas are called *Web sites*, the information display is provided by the experiment software and not a real Internet Web site.

- Factoids in your inbox can be copied into your MyFactoids list by selecting the factoid and clicking on the Add to MyFactoids menu item.
- To Post a factoid to a Web site that can be viewed by other participants in your group, select the factoid from either your inbox or your MyFactoids list that you wish to post, and click on the Post menu item, and in a popup window select the Web site you wish to post to.
- To Share, select the factoid from either your inbox or your MyFactoids list that you wish to share, click on the Share menu item, and select the anonymized name of the person with whom you want to share from a popup window. This sends the factoid to the selected participant's inbox.
- The Add to My Factoids menu item can also be used to copy a factoid from a Web site to your personal MyFactoids list. To copy a factoid from a Web site to your MyFactoids list, select the factoid you wish to copy from a Web site and click on the Add to MyFactoids menu item.

In summary:

- You can **Add** a factoid in your inbox to your MyFactoids list
- You can **Share** a factoid with another participant in your group
- You can **Post** a factoid to one of the "Web sites"

ICCRTS - ELICIT

- You can **Add** (copy) a factoid you find on one of the "Web sites" to your MyFactoids list.

You can also:

- Get a summary list of all the factoids in your MyFactoids list by clicking on the MyFactoids tab in the middle of your application window.
- Find out your role information and how other members of your group see you by clicking on the "How I'm seen" tab.
- Get a list of all the participants in your group, with information about their role and country, by clicking on the "What I see" tab.
- Access information from a Web site, by clicking on the Web site that you wish to view. To update the Web site with the latest information that has been posted to it, click on the Refresh menu item at the top of the screen, while viewing the Web site.

During each experiment round, you are free to work on any aspect of the task. The experiment software provides communications tools that you can use to interact with the other members of your group and communicate information.

When you think that you have identified the who, what, where and when of the adversary attack, click on the Identify menu item at the top of your screen and enter free text messages that identify the who, what, where and when of an adversary attack. Partial answers are accepted. You may **Identify** more than once.

- The *who* is a group (for example the blue group).
- The *what* is a type of target (for example an embassy or religious school or dignitary.)
- The *where* is the country in which the attack will take place (for example Alphaland.)
- The *when* is the month, day and hour on which the attack will occur (for example December 15, at 3:00 am.)

You can refer to these instructions during the course of the experiment by clicking again on the URL in the moderator message.

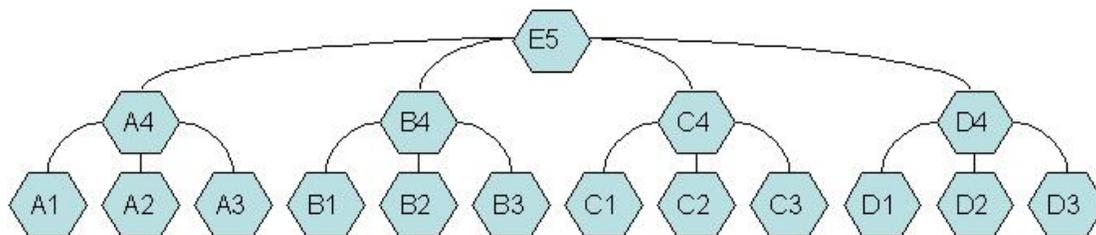
When you have finished reading this document, and are ready to begin the practice round, click the Ready menu item.

Appendix D: Instructions for C2 (B) group

"B" Group

You have been randomly assigned to the "B" group. During the trial you will be analyzing information factoids that are sent to you. You will also be able to **Share** information factoids with any participant of your group.

In the "B" group, there are four teams of four members each plus an overall cross-team coordinator. The four teams are organized along the lines of a traditional hierarchical command structure, each with a leader. This diagram shows the hierarchical relationship between a cross-team coordinator (E5), the four leaders (A4, B4, C4 and D4) and their team members:



Team A is focused on who, team B on what, team C on where and team D on when. The overall coordinator coordinates information between the team leaders across team boundaries. Note that the above diagram is an organization chart and not a communications chart.

The experiment software application supports two ways of informing participants in your group about factoids you have received. You can **Share** a factoid directly with another participant using the Share menu item or you can **Post** a factoid to your team Web site. Each of the four teams in your group has its own Web site.

Participants can post intelligence information (factoids) they have to a Web site that others in their team can simultaneously see and post to. Though these areas are called *Web sites*, the information display is provided by the experiment software and not a real Internet Web site. In the "B" group each team can see and update their team's Web site. For example, all members of the "where" team can see items posted

ICCRTS - ELICIT

to their team's "where" Web site. The cross-team coordinator can see all four of the typed Web sites

- Factoids in your inbox can be copied into your MyFactoids list by selecting the factoid and clicking on the Add to MyFactoids menu item.
- To Post a factoid to a Web site that can be viewed by other participants on your team, select the factoid from either your inbox or your MyFactoids list that you wish to post, and click on the Post menu item, and in a popup window select the Web site you wish to post to.
- To Share, select the factoid from either your inbox or your MyFactoids list that you wish to share, click on the Share menu item, and select the anonymized name of the person with whom you want to share from a popup window. This sends the factoid to the selected participant's inbox.
- The Add to My Factoids menu item can also be used to copy a factoid from a Web site to your MyFactoids list. To copy a factoid from a Web site to your personal MyFactoids list, select the factoid you wish to copy from a Web site and click on the Add to MyFactoids menu item.

In summary:

- You can **Add** a factoid in your inbox to your MyFactoids list
- You can **Share** a factoid with another participant in your group
- You can **Post** a factoid to one of the "Web sites"
- You can **Add** (copy) a factoid you find on one of the "Web sites" to your personal MyFactoids list.

You can also:

- Get a summary list of all the factoids in your MyFactoids list by clicking on the **MyFactoids** tab in the middle of the application window.
- Find out your role information and how other members of your group see you by clicking on the "**How I'm seen**" tab.
- Get a list of all the members in your group, with information about their role and country, by clicking on the "**What I see**" tab.
- Access information from a Web site by clicking on the Web site that you wish to view. To update the Web site with the latest information that has been posted to it, click on the Refresh menu item at the top of the screen, while viewing the Web site.

During each experiment round, you are free to work on any aspect of the task. The experiment software provides communications tools that you can use to interact with the other members of your group and communicate information.

ICCRTS - ELICIT

When you think that you have identified the who, what, where and when of the adversary attack, click on the Identify menu item at the top of your screen and enter free text messages that identify the who, what, where and when of an adversary attack. Partial answers are accepted. You may **Identify** more than once.

- The *who* is a group (for example the blue group).
- The *what* is a type of target (for example an embassy or religious school or dignitary.)
- The *where* is the country in which the attack will take place (for example Alphaland.)
- The *when* is the month, day and hour on which the attack will occur (for example December 15, at 3:00 am.)

You can refer to these instructions during the course of the experiment by clicking again on the URL in the moderator message.

When you have finished reading this important document, and are ready to begin the practice round, click the Ready menu item.

Appendix E: Experiment Procedures

An experiment consists of 34 subjects randomly assigned to the E or C2 variants (17 in each group). The C2 and E groups are run in different sessions. For each session, the subjects participate in two trials a practice round and an experiment round.

Below is an example sequence of events from the ELICIT welcome slide set.

- 1:00-1:10 Welcome Remarks
- 1:10-1:20 Pre-Briefing Video
- 1:20-1:50 Game I – Familiarization [No Talking]
- 1:50-2:05 – Break – [No Talking about the game]
- 2:15-3:15 Game II – [No Talking]
- 3:15-3:35 Game II Survey
- 3:35-3:40 Debriefing Video
- 3:40-4:00 Questions, Thoughts, Comments

A copy of these slides and the pre- and post-briefing videos and other experiment materials such as the moderator check list will be available at

<http://www.dodccrp.org/html3/elicit.html>

Before the experiment, the moderator plays a pre-briefing video that explains the experiment software and task.

A practice round allows the subjects to become familiar with the experiment task and software.

The experiment software automatically displays the URL for the appropriate survey at the end of the final round. The survey contains demographic questions, and a social networking survey instrument to manually measure trust, reputation and relationships.

After the survey has been completed, the moderator plays a post-experiment briefing.

Any additional questions the subjects have are answered. In classroom circumstances, a discussion of the issues at hand may then be held.

Appendix F: Experiment Software

The ELICIT software includes the server software and the client software.

Installation instructions for each of these applications are provided in Appendix G.

Technology Platform Requirements

The following computer resources are required to conduct the experiment program:

The ELICIT software runs on Windows XP with a minimum of 256MB, a local area network connection and (preferably) with SVGA (800x600) or higher resolution displays.

Each subject must have their own Windows XP system. The system must be running Java 1.4.2 or greater.

Moderators must be allowed to connect a Windows XP machine to the local network for use as the experiment server.

There must be a network Internet or intranet to support communications among the subjects.

Moderators must have sufficient access rights to install applications software on systems. This software is de-installed at the conclusion of the experiment.

The network must be unclassified.

As part of the process of installing the software for the experiment, the clocks (date/time) on the subject machines must be synchronized.

NOTE: A minimum of 18 machines are needed for each session. (17 subject machines plus one moderator machine). It is more convenient if there are separate moderator machines for the server software and for running the briefing videos, etc.).

Server Software

The server software (ELICITServer.msi) supports the interaction tasks among the subjects in the experiment. The server contains all the information needed to set up the experiment (name lists and country information, factoids, etc.) It also collects all experiment actions in the action log file (described in Appendix I) for each experiment trial.

The server software user interface allows the experiment moderator to specify the type of experiment to be run and the tables that are used to configure the experiment.

ICCRTS - ELICIT

When the server software is first launched, it prompts the moderator to specify the number of subjects in the experiment trial. This is generally 17. For test and demonstration purposes it could be 5 or even 1. The maximum number that has been run is 34, which is the number needed to run the E and C2 groups concurrently.

Experiments can be configured to run in trial sets of up to four trials each. Subjects are randomly assigned to roles at the beginning of a trial set.

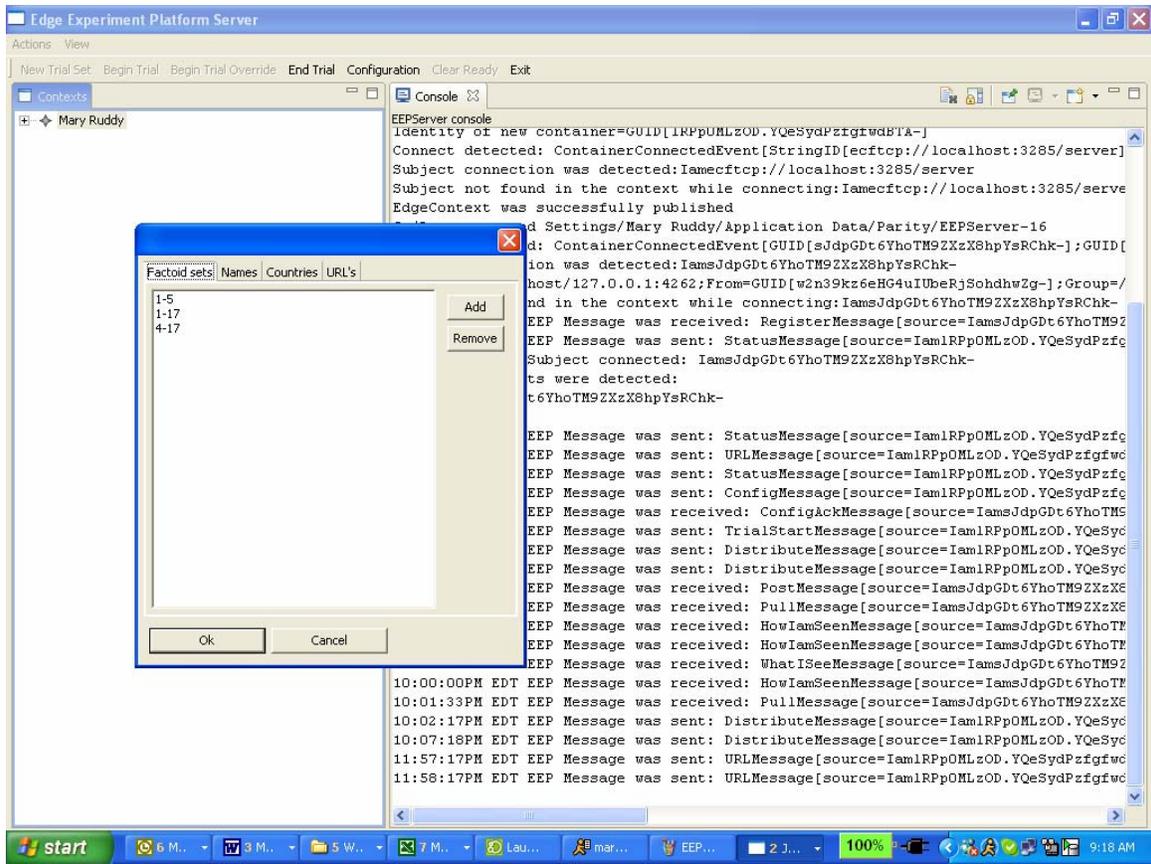
Actions

The server supports a number of actions that allow the experiment moderator to manage the experiment. Additionally a dashboard is available for monitoring experiments in progress.

Configuration

Before beginning an experiment trial set, the server must be configured. In anticipation of the need to run additional related experiments in the future, much more flexibility has been built into the server than has been utilized for the experiments run to date. Additional configuration flexibility is being added to the next release.

ICCRTS - ELICIT



The Factoid set to be used must be specified. See Appendix H for additional information on factoids and <http://www.dodccrp.org/html3/elicite.html> for a copy of all base factoid sets created.

The table that contains the subject names can be specified here. To date only one Names table has been created.

The Countries table must also be specified. The Countries table must match the Factoid set.

At specified times during an experiment, the URL's specified here are displayed to the subjects.

The first pair specifies the instruction URL that will be distributed to the subjects – one for the A (Edge) subjects and another for the B (C2 Hierarchy) subjects. This pair is required.

The second pair specifies the survey URL that is displayed at the end of each trial set. This pair is required.

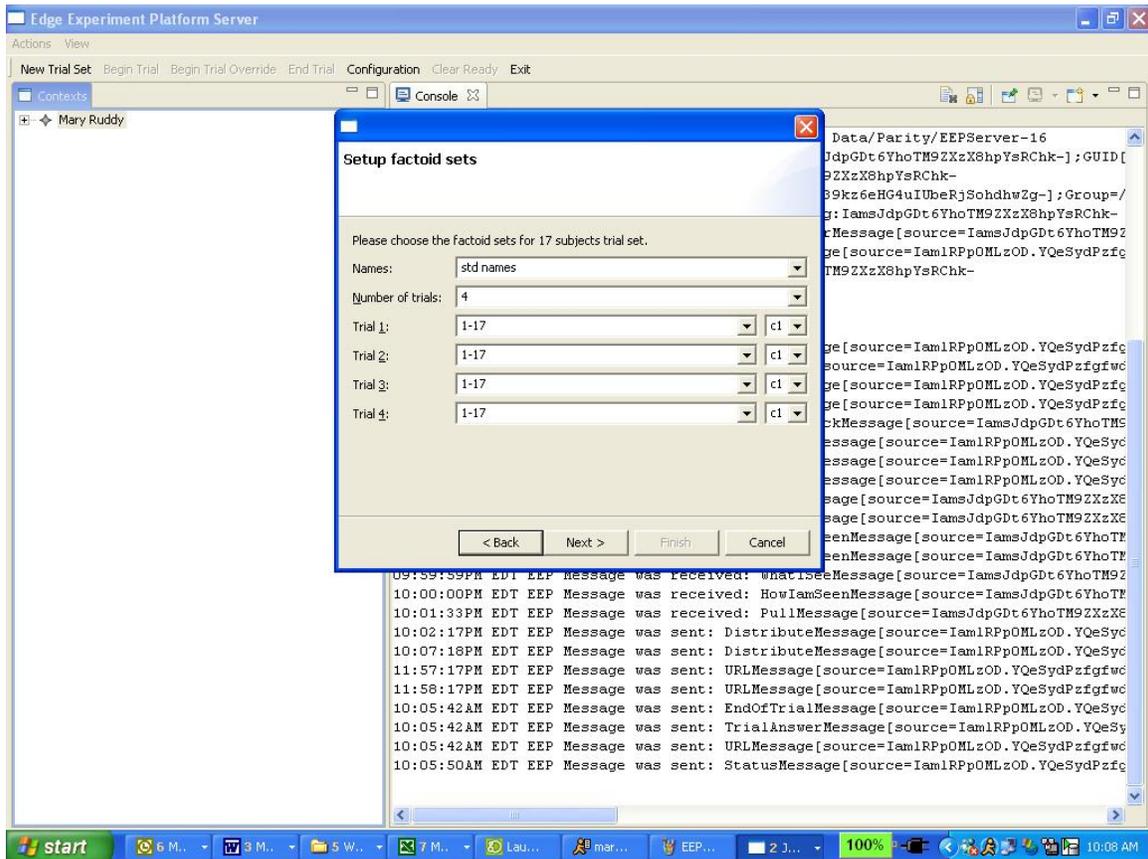
The other three pairs are optional. The first two of these can be configured to be displayed at a specified time during a trial round. Time is specified in minutes into the trial round. The final optional pair is displayed at the end of a trial set.

Beginning a Trial Set

Once the specified number of clients have connected to the server, a trial set can begin. For testing and demonstration purposes, it is possible to begin a trial set with fewer than the requested number of clients, by selecting the Begin Trial Override action.

The New Trial Set action allows you to specify how the trial set will be run. First it prompts for whether the trial set is for Edge community (A) or Hierarchy organization (B) or both.

It then prompts you to specify which of the configured data sets will be used in the trial set.



ICCRTS - ELICIT

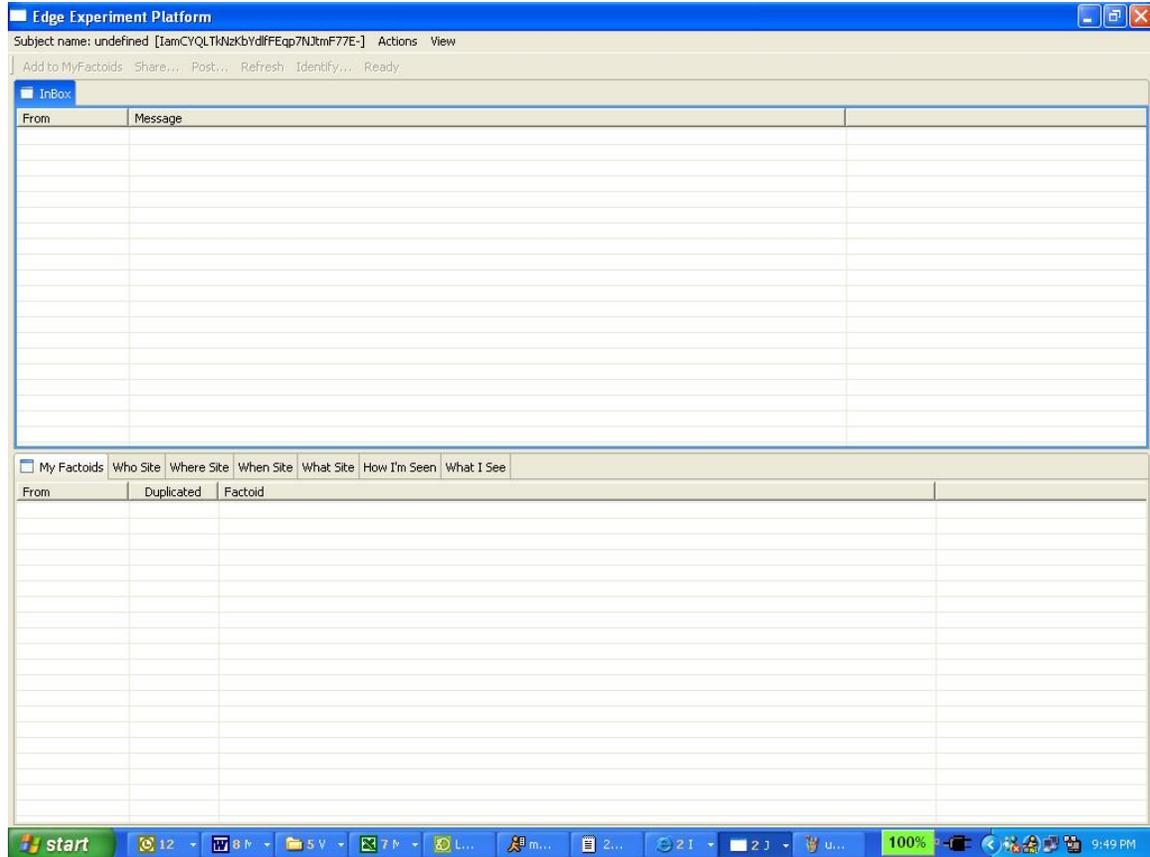
Finally the moderator is prompted to specify the label that subjects will see for messages that come from the server. By convention for these experiments, we have been using “Moderator-*<initials of moderator>*.”

Additional information about the operation of the experiment server software can be found in the moderator/coordinator checklist at <http://www.dodccrp.org/html3/elicite.html>.

Client Software

The client software (ELICITClient.msi) provides the user interface for ELICIT. It is the sole communication tool used by team members to obtain and share information needed to accomplish the assigned task.

The following illustration shows the main screen of a sample client user interface.



The upper part of the user interface contains the Action and View menu, as well as the InBox display.

The Actions menu includes the following items: Share, Post, Refresh and Identify

The View menu enables subjects to view the client console, which reports recent activity for testing and debugging purposes.

The InBox display contains a queue of messages received by the subject. Messages include both administrative information about the management of the experiment

ICCRTS - ELICIT

such as the beginning or end of experiment rounds, as well as any Share messages received from other subjects.

The lower part of the user interface contains a series of tabs that support the analysis and interaction tasks of the experiment. The subject clicks a tab to see the information displayed under that subject.

The MyFactoids tab lists the subject's own factoids (visible only to them).

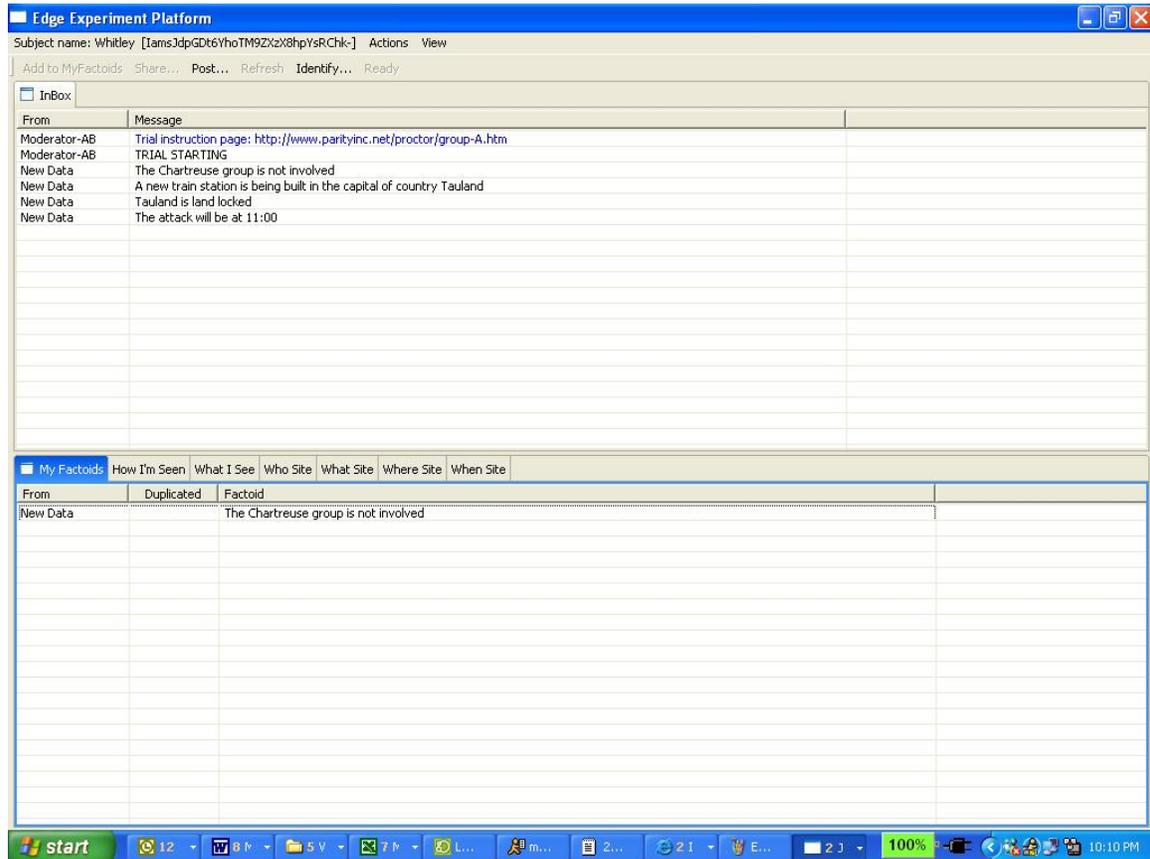
The Web site displays allow subjects to obtain information from available "Web sites" by clicking the available tab. The E community can access all four of the different Web sites (the Who site, What site, Where site or When site) for information about each aspect of the task. They must choose which aspect they wish to view by clicking a specific tab. The C2 subjects only have access to their own team's Web site. (e.g. either the Who site, What site, Where site or When site).

The How I'm Seen tab displays the subject's name, team assignment and in the C2 case, role.

The What I See tab shows the (anonymous) names of the 16 other subjects in the subject's experiment group. In the C2 case, the role assignment (team member, team leader or cross-team coordinator) is also displayed. When reputation metrics are displayed, they will be shown here.

The following sections illustrate portions of the user interface populated with data from an experiment trial.

Sample View of InBox

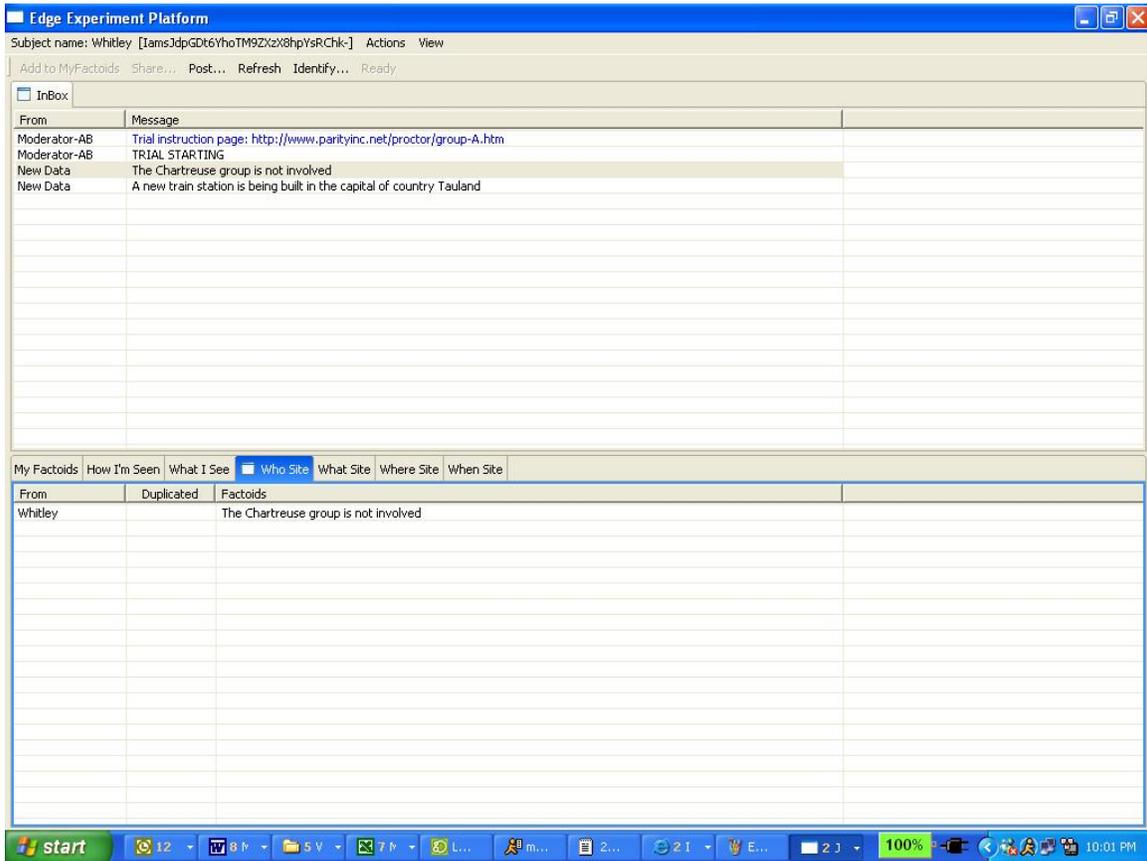


The InBox view at the top of the screen contains a scrolling list of messages that this subject has received.

Three times during a trial round each subject receives message(s) from the system, indicating that New Data has been provided to the subject.

NOTE: It is possible to configure the text of this New Data tag to perform additional experiments on the effect of the data source on subsequent data use.

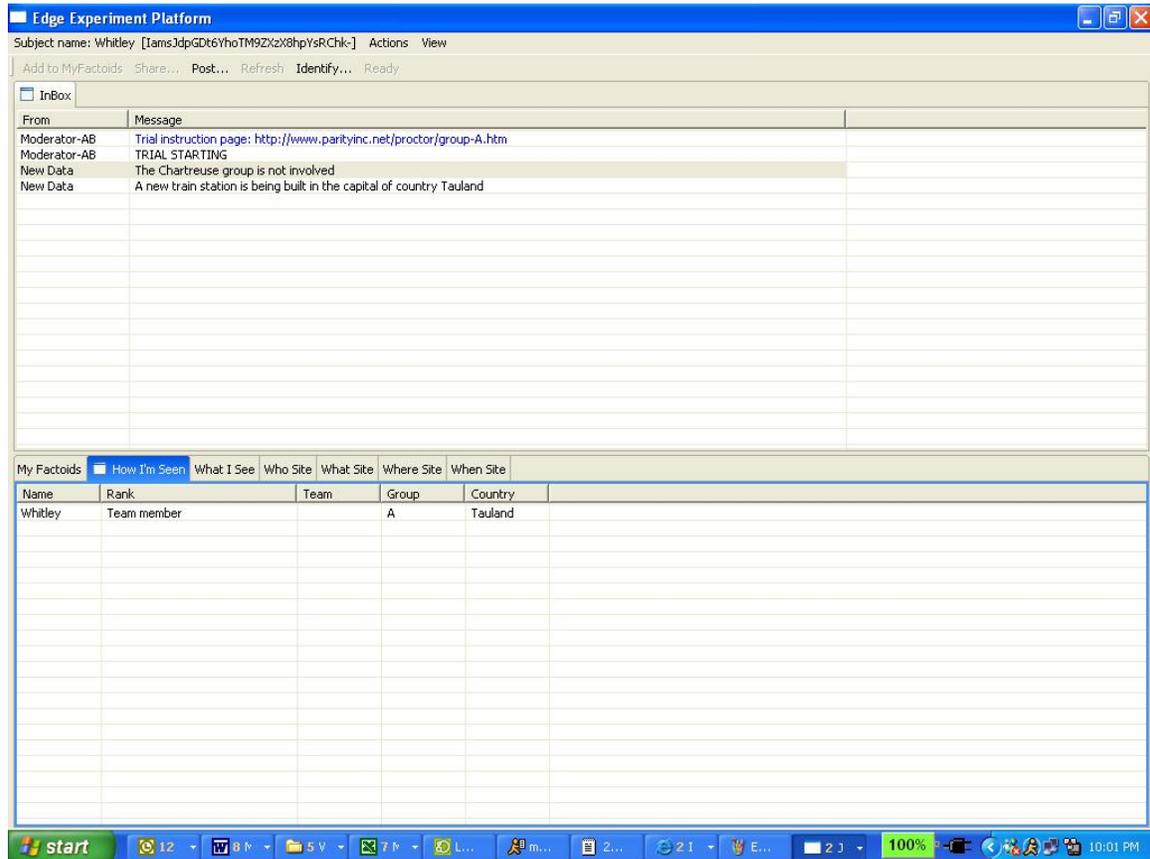
Sample View of Who Site



Web sites, like the Who Site shown here, contain a list of all factoids that have been posted to them by subjects.

The view in this example is the E variant, so four Web sites are shown (Who, Where When and What).

Sample View of How I'm Seen



NOTE: In the E variant, the Rank field is always Team member and the Team is not present. Possible values of the Rank field include Team Leader, Cross-team Coordinator, or Team Member.

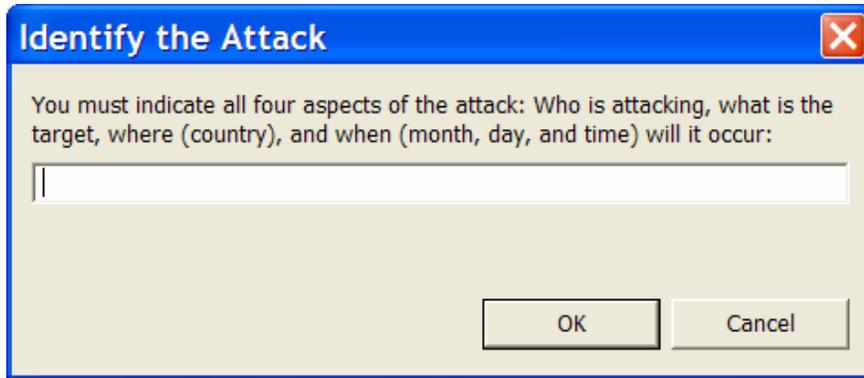
Sample View of What I See

The What I See view is similar to How I'm Seen, except that it provides information on the other 16 subjects.

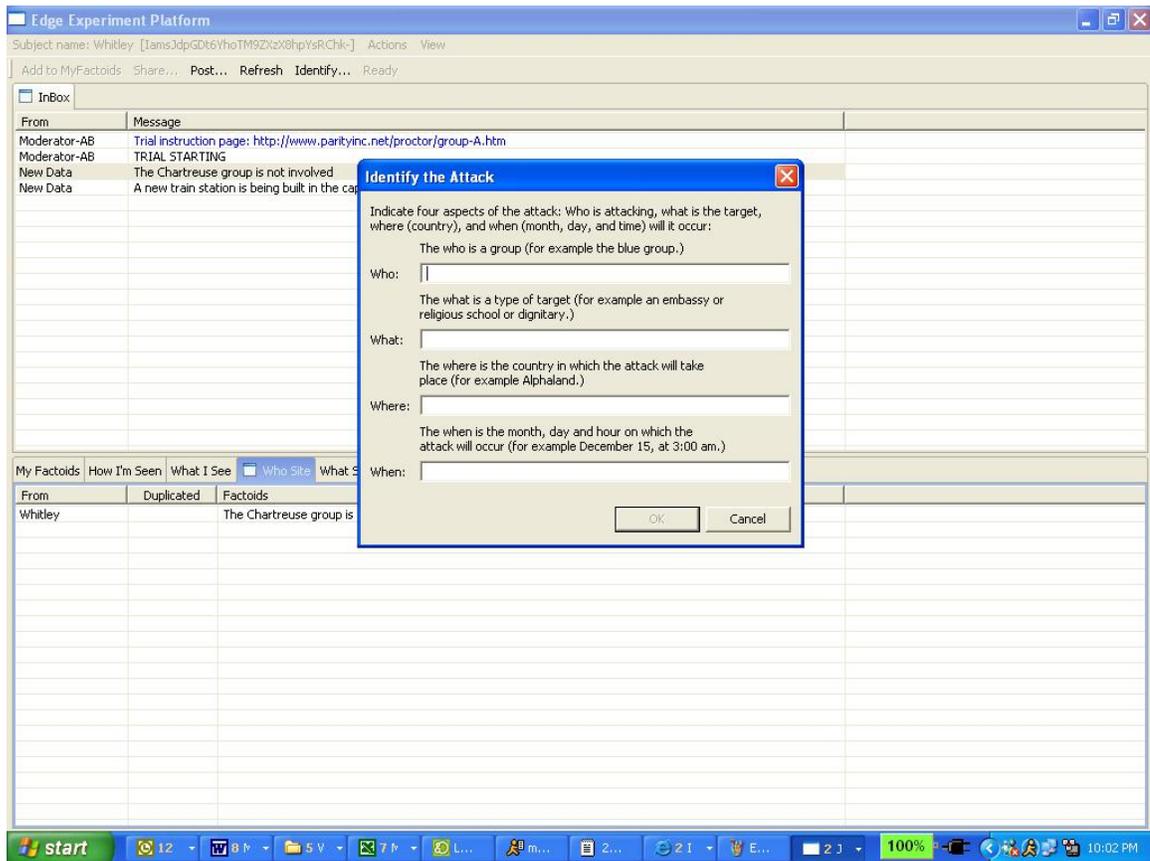
Sample View of Actions | Identify... Action

At any time during the experiments, subjects can click the Actions | Identify... action to open the Identify the Attack dialog box.

1. Subjects may type an arbitrarily long line of text into the text field.



2. When you click OK, the software sends the free text message to the experiment server.
3. Based on feedback received during the initial pre-test experiment runs at Evidence Based Research, this user interface was modified so that the identify information was entered in four parts:



Sample View of Actions | Share...

When subjects select a factoid from the My Factoids view or the InBox, they have an opportunity to forward the factoid to one or more members of their organization.

ICCRTS - ELICIT

1. The dialog box displays a list of eligible recipients with whom the factoid can be shared.
2. The subject can click OK to transmit the factoid to the chosen recipient(s)' InBox view or click Cancel to cancel the action.

Sample Views of Actions | Post...

If subjects select a factoid from the My Factoids view or the InBox view the results vary, depending on whether the subject is a member of the E community or the C2 community.

In the E-variant:

1. The dialog box displays a vertical list of four Web sites (who, what where and when) to which the factoid could be sent, with OK and Cancel buttons to complete the action.
2. If the subject clicks OK, the factoid is transmitted to the selected Web site.

In the C2-variant:

1. The dialog box displays a vertical list of one Web site (who, what where or when) to which the factoid could be sent, with OK and Cancel buttons to complete the action.
2. If the subject clicks OK, the factoid is transmitted to the selected Web site.

Appendix G: Installing The ELICIT Software

This appendix provides instructions for installing and de-installing the client and server applications on a single Windows XP machine for test or demonstration purposes.

1. Check what version of Java you have
 1. Point your browser at <http://www.java.com/> and click on "verify installation" and click the Verify Installation button on the following page. You should see something like: "java version "1.5.0_06"
 2. If you don't have Java or don't have at least version 1.4.2, then point your browser at <http://www.java.com/> and under "Free Java Software" click "Download Now"

Install the Server software

1. Uninstall the previous version of ELICITServer (the ELICIT Software) if any
 1. Click on Start > Control Panel >
 2. Choose "Add or remove programs"
 3. Search for ELICIT Server. If it is on the list, remove it
2. Download the ELICIT software
 1. Click on [ELICITServer.msi](#) (14.5MB Windows .MSI installer download), choose "Save." Save it on your desktop and click "Save"
3. Install the ELICIT software
 1. Double click on the downloaded ELICITServer.msi file on your desktop to install the software.

Install the Client software

1. Uninstall the previous version of ELICIT (the ELICIT Client Software) if any
 1. Click on Start > Control Panel >
 2. Choose "Add or remove programs"
 3. Search for ELICIT. If it is on the list, remove it
2. Download the ELICIT software
 1. Click on [ELICITClient.msi](#) (14.5MB Windows .MSI installer download), choose "Save." Save it on your desktop and click "Save"
3. Install the ELICIT software
 1. Double click on the download ELICITClient.msi. file on your desktop to install the software. Once you finish the installation wizard, a desktop icon named "ELICIT: will appear.

Set machine configurations (setting Always on is not necessary for casual tests)

1. Configure your machine so that it is "Always On"
 1. Click on Start > Control Panel
 2. Click on Performance and Maintenance
 3. Click on Power Options
 4. On the tab marked, "Power Schemes," click on the drop-down menu, choose "Always On" and click "OK" to close the window.
2. Configure your machine so that all machines are using the same time zone. (e.g. Eastern Daylight Time.)
 1. Click on Start > Control Panel
 2. Click on Date, Time, Language and Regional Options
 3. Click on Date And Time
 4. It should say your time zone time. If it doesn't, click on the Time Zone tab and select your time zone time from the drop down and check the box for daylight savings time
 5. Click "OK" to close the window
3. Start the server
 1. Click on Start >
 2. Click on All Programs
 3. Click on Parity
 4. Click on ELICIT server
 5. Select the number of clients you will be using, (1 or 5 recommended for preliminary tests)
4. Start the clients(s)
5. Double-click on your ELICIT desktop icon (A blue sphere with stripes labeled ELICIT)
 1. If the software prompts you for an ELICIT server address, you are all set. Enter the IP address of the server or localhost if the client and server are on the same machine.
 2. You can delete ELICITClient.msi from the desktop
6. Call for instructions about configuring and running the game or use the moderator coordinator check list as a start.

To de-install the experiment software

1. Click Start.
2. Click Control Panel.
3. Click Add/Remove Programs.
4. Select ELICIT***.
5. Click Remove.

Appendix H: Factoid Sets and Factoid Keys

This appendix provides detailed information about the structure of factoid sets used in the experiment, as well as the structure of the factoid keys used to identify the data. Each factoid is a logical statement which must be combined with other factoids to deduce the result.

Some factoids are phrased as positive statements others are negative. Positive factoids may be either existential or universal. For example, if the following three relevant factoids (one positive/existential, two negative) were available, the location can be determined correctly as London:

1. It might be in London, New York, Paris, or Ottawa
2. It will not be in North America
3. It will not be in Paris nor Tokyo

In this example, the first factoid indicates four *possible* locations, although it is also possible that the incident will occur in Moscow or Lima or somewhere else. An example of a universal positive factoid would be phrased: “It will be in London or Paris.”

Negative factoids are definite statements that are universally true within the scope of the experiment.

Order Dependence

Some factoids can only be partially or fully applied after other conclusions have been reached. For example, the factoid “The city’s name begins with an “L” is only useful after one or more items are listed that start with that letter.

Completeness and Redundancy

The factoid set is designed such that if every factoid is used, the correct conclusion is reached. However it is also true that not all the factoids are needed to correctly identify the adversary attack. As in the real world, some of the factoids are key or expertise factoids (25%), some are supportive (25%) and many are just noise (50%).

All subjects will continue until the ELICIT software (as controlled by the experiment’s moderator) declares that the task has been completed.

Factoid Sets Created

Four factoid sets were created as part of this project. The factoids were intentionally constructed so that they were evenly distributed over a number of categories.

Factoid set 4 is the easiest. Factoid set 1 is next in degree of difficulty. Factoid set 2 and 3 are very similar. The names used in constructing the factoid sets were abstracted so as not to evoke any current political bias:

- Adversary groups were named for colors.
- Countries were named for Greek letters.

Each subject receives four factoids:

- One that is key or expert information.
- One that is supportive.
- Two that are noise.

Factoid sets are stored as tables and include the answer to the task. Factoid sets were tested to remove unintended ambiguity and to ensure that persons presented with all the factoids would come to the same conclusion about the who, what, where and when of the adversary attack.

Factoid sets were also designed so that they could be modified or additional factoid sets created without requiring any code changes. For planning purposes, the amount of time that would be needed to create new versions of the factoid sets that were logically equivalent, yet used different names, is minimal. Creating significantly new factoid sets that work with the existing software architecture (i.e., that don't require code changes) yet have varying internal logic was found to require several days of work. This is due to the constraints introduced by

- Ensuring that each subject receives factoids of equivalent value.
- Multiple rounds of testing required for the modified factoid sets.

Naming Conventions

Factoid set names have two components. The first specifies the scenario that is being used (1..4). The second specifies the number of subjects over which these factoids are distributed. Seventeen is the number used for actual experiments. For testing

ICCRTS - ELICIT

and demonstration purposes, the numbers 1 and 5 are also used. Thus 1-17 and 4-17 are valid factoid set names for experiment runs.

Solution Requirements

Copies of the factoid sets are provided with the experiment materials. For this project, factoid set 1 was used for the experiment trials.

In factoid set 1, the tasks of identifying the who, what, where and are of varying difficulty. For example, identifying:

- Who required 5 factoids to formally solve
- What required 9 factoids
- Where required 9 factoids (the same nine used in What)
- When required:
 - 1 factoids to identify day
 - 2 factoids to identify month
 - 1 factoids to identify hour
 - 1 factoid to identify AM/PM

Different factoids required differing levels of logical effort. The logic required to analyze:

- Who factoids to determine the Who is moderate
- What factoids to determine the What is difficult
- Where factoids to determine the Where is difficult
- When factoids to determine the When is trivial.

See the factoid solution requirements spreadsheet provided with the experiment materials which are available at <http://www.dodccrp.org/html3/elicite.html>

for a complete list of base factoids.

Some of the supporting factoids might be interpreted as substitutes for some of the key factoids. Also if a subject ignores the issue that Aphaland's embassies are not in Alphaland, but rather in host countries, it is possible to get the correct answer to What using fewer factoids. Therefore some of the identify questions may be easier than anticipated. Scoring of the difficulty of the identify subquestions is an area for further analysis. This ambiguity, and the fact that there are a finite number of

ICCRTS - ELICIT

possible answers, explains why sometimes people get the right answer when they do not have all the factoids to technically answer the question.

The names of the countries that subjects are assigned to are also stored in a table as are the names used to anonymize the subjects. For this project subject names are gender neutral and of approximately the same length, and are names that are commonly known to English-speaking audiences. For future experiments it would be possible to create additional name tables containing names that are familiar in other parts of the world

It would also be possible to create new versions of factoid sets that experiment with other factoid distributions. For example, in current factoid sets, factoids are distributed evenly (e.g., four to each subject). This could be changed so that some subjects get more than others). An uneven distribution would allocate more factoids to certain fortunate subjects at the expense of other less fortunate subjects.

Factoid Keys

The factoid keys are both unique within a factoid set and designed to describe the type of factoid they represent. The factoid key can be used to facilitate data analysis, by grouping factoids and the actions taken on them.

The factoid key contains six parts:

- Factoid set number [01 is the first factoid set]
- Factoid number in the factoid set (1...68)
- Distribution wave (1, 2 or 3)
- Type [1 is who, 2 is what, 3 is where and 4 is when]. If a factoid could be interpreted as applying to more than one type, the primary type is specified.
- Impact [E, K, N, S] The impact of the factoid.
 - E factoids represent Expert information.
 - K factoids represent Key information.
 - N factoids are noise. They are not important. The game can be correctly solved if these are ignored.
 - S factoids are supportive. These factoids support the information in the E and K factoids.
- The count of the factoid within its type-impact category. This count makes the factoid keys unique. For example, the 015732K1 key is:

ICCRTS - ELICIT

- from the 01 factoid set
- the 57th factoid in the sets
- distributed in the 3rd wave
- a “what” factoid
- a key factoid
- the first key factoid in the list of “what” factoids.

Appendix I: ELICIT Action Log

This appendix describes the action logs created by the ELICIT experiment trials.

Overview

One of the design goals of the experiment is to log and analyze all the actions of the experiment subjects. Because the analysis and interpretation of a subject's actions depends on their role in the experiment and the factoids they have been assigned, the log records all these assignments in the action log. The information is intentionally redundant to make analysis of isolated actions more efficient.

All entries in the log are date and time stamped.

Log Generation

A separate log file is created for each trial. (A set consisting of three trials would generate three separate log files.) The log files are date and time stamped with a file name in the following format:

YearMonthDay-HourMinute-xxxxx-Trial number in trial set

This number is followed by A for an Edge-only trial, B for a Command and Control-only trial or both A and B for trials that have both E and C2 groups. The xxxxx is a random number that is included for uniqueness.

The log file describes the characteristics of each experiment trial and correlates the trial type to the trial number. For example, a log named

20050629-1031-40757-2A.log

was created on June 29, 2005 at 10:31 AM for the second trial in a trial set for an Edge group.

Actions

Log entries are organized in lines. Each entry consists of fields that are delimited by spaces.

Each entry contains the following fields:

ICCRTS - ELICIT

- Date in day/month/year format
- Time in hours:minutes:seconds
- Total number of seconds since the trial was initiated
- Entry type.

There are fifteen types of entries, which can be divided into two categories: Actions taken by the system and actions taken by a subject. The system actions are primarily for documentation purposes. The subject actions are intended to be imported into an Excel spreadsheet or other tool for detailed analysis.

- System action entries are written each time the system performs an action, such as:
 - initiate
 - factoid
 - context
 - facet
 - assign
 - admin (administration)
 - start
 - dist (distribution)
 - end.
- Subject action entries are written each time a subject performs an action, such as:
 - how_seen
 - what_see
 - share
 - pull
 - post
 - identity.

For example the line beginning,

8/05/2006 11:04:30 0000 – initiate

was written on August 5, 2006 at 11:04 and 30 seconds AM, when the trial was initiated (in the zeroth second of the trial.)

Each entry type also contains additional fields that vary by entry type. The following sections describe additional fields written for system and subject entry types.

System Action Entries

This section lists and describes the system action entries.

- initiate – the directory where the log file is being written and the file name of the factoid set being used.

- factoid – the factoid key and factoid text in quotations.

This provides a self-documenting key to all the factoids used in the trial. These entries duplicate the information in the factoid set so that it is readily available in the log.

- context – type of trial (A for Edge community or B for C2 organization), moderator name entered into the server dashboard, and number of contexts.

A context is a system resource. The server is a context and each client system is a context.

- facet – internal key assigned in a context.

This key is also displayed at the top of each subject's application right after their name. The facet key is used for testing purposes. It is not expected to be used as part of the post experiment analysis.

- assign – facet key, subject assigned name, subject group, subject team, subject role, subject country.

This documents which subject was assigned to what role. Thus we know, for example, that in this round "Chris" was the leader of the "Who" team in the C2 group in Chiland.

- admin (administration) – facet key, subject assigned name, subject group, subject team and either CONFIG for configuration, CONFIG ACK for configuration acknowledgement, START to confirm that the trial has started for that subject or trial survey page.

Since team synchronization is an important element of the experiment, we log each time the server communicates with a subject/client. At the beginning of each trial set, the server sends a message to each subject machine and waits until each message has been acknowledged before proceeding to actually start a trial.

Once a trial starts, another set of messages is sent to each subject indicating that the trial has started and providing a URL for instructions.

At the end of trial, a survey page URL is sent to each subject. Up to three additional optional surveys can also be sent during a trial.

- start – indicates that a trial has actually started. Includes the trial number (1 to 4).
- dist (distribution) – the system fully documents which factoids were distributed to which subjects.

Factoids are randomly distributed to subjects three times during each trial: at the beginning, at the 5-minute mark and at the 10-minute mark. Each time a factoid is distributed, a log entry records the factoid key of the subject, the subject's assigned name, the subject's assigned group, the team they are on, the key of the factoid distributed and the text of the factoid in quotations.

- end – indicates that a trial has actually ended. Includes the trial number (1 to 4).

Subject Action Entries

This section lists and describes the subject action entries.

- **how_seen**

Any time a subject clicks the “How I’m Seen” tab to see their assigned role, a how_seen entry is written. This entry includes the context key, assigned name, group and team of the subject requesting the “How I’m Seen” information.

- **what_see**

Any time a subject clicks the “What I See” tab to see information about their group, a what_see [sic] entry is written. This entry includes the context key, assigned name, group and team of the subject requesting the “What I See” information.

- **share**

Any time a subject shares a factoid with another subject, a share entry is written. The entry includes the context key, assigned name, group and team of the subject initiating the share; the factoid key and text; and the context key, assigned name, group and team of the subject receiving the shared item.

- **pull**

Any time a subject clicks the tab of a Who, What, Where or When site to see its contents, a pull entry is written. The entry includes the context key, assigned name, group, and team of the subject initiating the pull and the name of the site being pulled.

- **post**

Any time a subject posts a factoid to a Who, What, Where or When site, a post entry is written. The entry includes the context key, assigned name, group and team of the subject initiating the Post; the factoid key and text, and the name of the site to which the item is being posted.

- **identify**

When a subject provides information on the Who, What, Where or When of an adversary attack, an identify entry is written. The entry includes the context key, assigned name, group and team of the subject initiating the identify; followed by the Who, What, Where and When information strings.

- end – indicates that a trial has ended. Includes the trial number (1 to 4).

Appendix J: Creating Data Analysis Files Post-Experiment Analysis

The experiment server software creates a log that records all actions taken by the experiment's server and every subject in the trial. For each experiment run, the server log, together with the Web-based survey data is available at <http://www.dodccrp.org/html3/elicite.html>. Subjects and the Moderator /coordinator checklists are also available at this Web site.

Including pre-test runs, there have been 10 live subject runs with 17 subjects each using the standard experiment protocol. In addition there was a pre-alpha test with 5 subjects. A complete list of these runs and the data files associated with them, metadata and comments is contained in Summary of Experiment Trials v2.xls which is available at <http://www.dodccrp.org/html3/elicite.html>. Though most of these runs are not suitable for full analysis, they are all being provided as a complete record of all experience to date. It is anticipated that future projects may wish to do further analysis in the area of self synchronization, even in situations where it is not possible to fully measure task effectiveness.

Instructions for creating a set of Excel spreadsheets for intermediate analysis from these log files are provided in this appendix. Using the methodology outlined below, the raw data is extracted and sorted to make it easier to answer such questions as “did the subject reciprocate after receiving a shared factoid”, or “what was their shared awareness at a particular time.” The list of analysis workbooks created as part of this project together with all of the workbooks is available at <http://www.dodccrp.org/html3/elicite.html>

Because so few people completely identified the adversary attack, we did not perform statistical analysis on the length of time it took to identify the adversary attack. See the Preliminary Results and Conclusions section and Appendix L, Suggestions for Future Work sections for recommendations of additional efforts.

Note that throughout this project we continued to evolve the experiment platform and procedures. This included for example fixing a factoid set distribution error after the May 16 pre-test before the live subject tests performed on May 27. (One factoid was going to the wrong person, causing one person to receive an extra factoid and another person to receive one fewer factoid.)

ICCRTS - ELICIT

This appendix describes a procedure for creating Excel files from the raw data log files generated in the experiment trials. (Note: another approach is being developed by a Canadian research team.) The appendix provides instructions for creating the following workbook files:

- Data – Takes raw data and organizes it
- Subjects – Shows a record of the subject's actions for the experiment
- Posts – Shows information available on a posting board
- Compare – Shows what clues a subject could see when
- Identify – Shows conclusions the subjects reached.

Creating the Data File (dateData.xls)

1. Open a new workbook and save it as *dateData.xls*. Use the Import External Data command on the Data menu to import the data into an Excel workbook using the defaults. Raw log files is available at <http://www.dodccrp.org/html3/elicite.html>
2. Insert a second worksheet into the workbook file and name it Mod Data. Copy the data into the Mod Data worksheet.
3. Delete unnecessary columns (A [date], B [time in hours], D [dash], F [computer reference]).
4. Insert a row at the top for the column headings. Label the columns Time (sec) [A], Action [B], Action By [C], Information [D] and Sent/Posted To [E]. Sort by Action.
5. Delete the admin rows, context row, end row, facet rows, factoid rows, initiate row, and start row.
6. Select all and set the font to Courier.
7. Select all the cells in the Action By column and use the Text to Columns (TTC) command on the Data menu. Choose Fixed width and place a break after the 22-character computer code and dash. Make sure there are no other breaks in the lines.
8. Delete all the information currently stored in the Action By column and move the lines in the Information column to the Action By column.
9. For the rows with an action of "assign", separate the name from the other information using TTC. The name should be in Action By and the rest in Information.

NOTE: The TTC function also has the delimited option. If using this option, be sure to check the separation of data and make sure all of the data is in the appropriate column.

ICCRTS - ELICIT

10. As in step 9, use TTC to separate the data for the **dist** actions. With this set, delete the numbers that separate the subject's name and the clue, so the subject's name should be in Action By and the clue in Information..
11. Leave the **how_seen** actions alone and move to the **identify** actions. Using TTC, separate the name from the response.

NOTE: Using the delimit option works well here using the “ (double quote) symbol as a separator.

12. For the **post** action, use TTC to separate the name, 8-character computer codes, clue and posting place. Delete the computer code. The name should be in Action By, the clue in Information and the posting place in Sent/Posted To.
13. For the **pull** action, use TTC to separate the name and posting board accessed (who, what, etc.). The board accessed should be in Information.
14. For the **share** action, use TTC to separate the name, 8-character computer code, clue, 22-character computer code and dash, and the recipient of the share. Delete the two computer codes. The names should be in Action By, clues in Information and the recipient of the **share** in Sent/Posted To.
15. Leave the **what_see** data alone and return to the top of sheet. Check again to make sure the data that was separated is still in the appropriate column and has not been moved while separating other sections of information.
16. Select the cells in the header and use the AutoFilter option under Data. This creates drop-down menus that contain various types of information. Check these menus to confirm that every cell is in the appropriate place.
17. Save the workbook.

Creating the Subjects File (dateSubjects.xls)

1. Open the organized data file created from the raw data (*dateData.xls*).
2. Create a new workbook called *DateSubjects.xls*. In this workbook, create 17 sheets, each named after one subject name in the experiment (e.g., Alex, Chris, Dale). Add the following column headings to each sheet: Name, Location, Team and Status.
3. Open the *DateData* workbook and use the drop-down menu in the Action By column to sort the information *in ascending order*.

NOTE: Do not overlook this important step.

4. Now select one name. Copy all the lines that are beneath the header. Paste these lines a few lines below the header to the appropriate tab. (Make sure you leave a few lines for header information and that the information is pasted in the appropriate columns.)
5. Take the assign line and fill out the information in the header. (Ignore the number if it is still part of the line.)
6. Return to the Data workbook. Set the Action By field to **All**. Go to the Sent/Posted drop-down column. Sort this information in ascending order.
7. Select the same name used in step 4. Copy the information underneath the header. Paste these lines immediately below the ones pasted in step 4. Change all of the shared Actions to **rcv_share**.
8. Create the same header as in the Data workbook, including the AutoFilter.
9. Repeat steps 3-8 for all 17 subjects.
10. Save the workbook.

Creating the Posts Data File (datePosts.xls)

1. Open the organized data file created from the raw data. (*dateData.xls*)
2. Open a new workbook and call it *datePosts.xls*. Add four worksheets to the workbook and label them Who, What, Where and When.
3. Go to the *DateData* workbook. Use the Sent/Posted To column to sort the information. Select one of the categories.
4. Copy the Time and Information columns. Paste them to the appropriate worksheet in the Posts workbook.
5. Use the Sort option to alphabetize the clues. Delete any posts that appear more than once, taking care to retain the one with the earliest time.
6. Repeat steps 3-5 for the other categories.
7. Save the workbook.

Creating the Compare File (dateCompare.xls)

1. Open the Subjects and Posts workbooks. Open a new workbook and call it *dateCompare.xls*.
2. Set up this workbook like the Subjects workbook, including the headings for name and such.

NOTE: You may prefer to copy the entire Subjects workbook, delete all the information under the headings and rename the file.

3. Copy the information from the Posts workbook that the Subjects would have access to.
 - a. In experiments that involve team and leaders (C2), subjects can only see the posting board of their team – except the cross-coordinator, who has the ability to see everything.
 - b. If there is no set structure to the group and everyone is a team member, everyone has access to every board.
4. Open the Subjects workbook. Sort the information by the Action column. Copy the time and information of the lines that are **dist** and **rcv_share** (the clues that the individual was sent). Paste these lines into the appropriate tab in the Compare workbook.
5. Alphabetize the clues. Delete any duplicate clues, taking care to retain the one with the earliest time
6. Save the workbook.

Creating the Identify File (dateIdentify.xls)

1. Open the Compare workbook. Compile or obtain the list of clues needed to solve the problem, and identify what part of the problem they are relevant to (who, what, etc.).
2. Copy this list of clues into one of the Compare worksheets. For later reference, create a table to mark the needed clues that the subject had access to and mark the clues that were on the page.
3. Complete step 2 for all tabs (17 total).
4. Open a new workbook and call it *dateIdentify.xls*. Open the Data workbook. In the Data workbook, sort the information by the Action column. Select the **identify** action. Copy the rows into an empty worksheet in the Identify workbook and name it Data.
5. Using the Text to Columns option, separate the data into separate columns. Delete any extraneous data.

NOTE: Using the delimit option with semicolons and equal signs works well.

ICCRTS - ELICIT

6. Create a table with the headings Subject [A], (blank) [B], Who [C], Person [D] What [E], Where [F], Month [G], Day [H], Hour [I] and AM/PM [J].
7. List the subject Names in alphabetical order, leaving a blank row between names.
8. In the cell to the right of each name, type “could solve.” Beneath that, type “did solve.”
9. Use the information located in the Compare workbook and fill in the **could solve** row.

NOTE: If one necessary clue from a section is missing, the section is considered unsolvable.

10. Go to the Data worksheet in the Identify workbook. Sort the rows with the subject name as the primary sort and time as the secondary.
11. Fill in the **did solve** row by looking at the data just sorted.
 - a. If the last answer is correct, the section is correct. Mark the cell with an X.
 - b. If the last answer is not correct, but the correct answer is present at an earlier time, mark the cell with a C [change].
 - c. If the correct line is not present anywhere, but an answer was given, mark the cell with a W.
 - d. If the no answer is given at all, mark the cell with a B.
12. Save the workbook.

Appendix K: Running the Experiments in a University Environment

This appendix provides pointers on some practical details of administering a series of ad hoc experiments at a private university.

Loading Software

It is recommended that subjects pre-install the experiment software on their computers prior to the start of the experiment. Though all won't pre-install, most will. This reduces the amount of time and installation questions that take place on site.

Hardware

Most (not all) undergraduate students interested in participating could arrange to bring a laptop. Some had to borrow a laptop. The quality of the laptops and their speed varied widely (20:1 ratio of time required to start the client application). We had three loaner machines and consistently used them all. We routinely had two people who didn't have machines to bring and pre-arranged use of a loaner machine. We also had a student or two who brought machines that were problematic and needed to be switched out. In the future it is recommended that we have more spare machines. In other words, required *all* students to bring their own machines, so that our machines are not all pre-allocated and can serve as actual spares.

Many laptops did not have Wi-Fi or had "flaky" Wi-Fi. A 24-port switch and 17 cables were purchased and brought to the experiment locations to ensure that we would be able to achieve connectivity. Typically we had to cable about 10 of the machines.

Timing

In general, if students are given the opportunity to pre-load the software, it takes about a half hour to get all laptops set up and ready to begin. Most are set to begin after about 15 minutes. Unfortunately, we frequently had to swap out machines after we thought everyone was all set, when a particular machine was not

ICCRTS - ELICIT

functioning well; this significantly extended the set-up period—often an additional 45 minutes.

In hindsight, having five Wi-Fi enabled replacement laptops would have saved time and frustration. A casual check on sale prices at Best Buy found laptops in the \$500 range. It is very inefficient to have a room full of people and realize that you really need one more mouse or laptop that works. Total elapsed time could have been noticeably reduced by having additional spare hardware.

The Boston University experiment runs were done with 20-minute practice rounds and a 60-minute experiment round. A full hour is too long for most subjects. We recommend reducing the experiment round to 45 minutes. The subjects became noticeably restless about 30 minutes into the experiment runs.

Twenty minutes is enough time for everyone to complete the survey.

We needed to be in the room a half hour to 1 hour before students arrived. Students were in the room for 3 ½ to 4 ½ hours – not including discussion. It took about another half hour to pack up.

Boredom

Since some people are much faster than others, future experimenters might consider suggesting that subjects bring a book to read in case they finish early.

Frequently Asked Questions

By software version 1.13 the user interface was so improved that questions about functionality virtually stopped. The moderator/coordinator checklist has been successively annotated to anticipate most of the other questions. The result is that we now have very few questions during the session after the initial hardware/software setup is complete.

Proctor Staffing

With current software and processes, two proctors/tech support people are sufficient for running an experiment in a single room. With the replacement laptops mentioned above, two people should also be sufficient for initial setup.

Recruiting Students

Different tactics need to be used during different parts of the school year.

Previously we had arranged to run the experiments in a classroom as part of a course during the school year. Some university (Non-DOD) experiments rounds were conducted after semester's end so it was necessary to find another source of students.

We used a student coordinator to find the student subjects. Recruiting the 17 subjects for an experiment round averaged about 6 hours of effort.

We paid the students a flat rate of \$100. This worked out to about \$20-25 per hour. Feedback is that the factors that influenced participation were (highest influence first):

1. The money
2. Participating because their friends were doing it
3. Playing the game

Any particular student coordinator has a finite circle of friends and organizations that they can leverage. This means that the more experiment rounds they staff, the more difficult it is for them to recruit people. This phenomenon was somewhat offset by the positive buzz generated by the experiment. The positive buzz to quote the coordinator was that it "was semi-interesting work for a short amount of time that pays really well."

The student coordinator leveraged their existing social ties: fraternities, several BU summer groups, and an athletic team.

The feedback from the student coordinator was that \$100 was a little high. In other words, it was good pay and caught the students' attention. He said:

"Offering a flat \$100 seemed strong and made the would-be test subjects take the offer very seriously and immediately volunteer their time. It would have taken longer had anything under \$75 been offered. It is obviously impossible to tell for sure. My initial reaction to the payment of \$100 was that it was a little high. "

ICCRTS - ELICIT

The summer is a very difficult time to recruit. The end of a semester and the beginning of the summer are particularly challenging times. It was much easier to recruit students during the academic year.

Appendix L: Suggestions for Future Work

Suggestions for future work can be grouped into three categories: software, experiments and analysis.

Software

The “client” user interface (used by the subjects) has already gone through several design iterations and seems to work well in well. We are in the process of creating a Web-based version of ELICIT to make it even easier to conduct additional experiments.

Now that we have several runs with live subject data and have more experience analyzing it, it is clear that further work could be done to automate the analysis process.

Experiment

Further refinement of hypotheses and better definitions

Although more analysis on the current data sets can be done to attempt to disprove the five Nulls stated in this report, it is also clear that we’ve learned much about what questions we should be asking (i.e., what new Nulls we should be attempting to disprove) and the need for better definitions of our terms.

Let us start with H1 as an example. Null hypothesis #1 states: “The platform-derived measures of trust and shared awareness do not correlate with the results of these same measures using traditional social network analysis of survey-gathered data.” This is fuzzy in a number of ways. First, the phrase “platform-derived measures of trust and shared awareness” is a composite of multiple sub-measures some of which do, and some of which do not, correlate with the survey-gathered data. The only way to make sense of this is to unpack the composite measure. We must propose separate nulls, each dealing with the correlation between platform- and survey-based methods for a separate sub-part of “trust and shared awareness.”

ICCRTS - ELICIT

Even “trust” needs to be unpacked a bit. For example, in the survey we found that, measured by the number of responses to questions dealing with who you would, or would not want in your group in the future, there were more positive responses in the Edge trials than in the Hierarchy trials. This would indicate more trust in the Edge group. To contradict this finding, we can look at different aspect of trust. If we look at the combination of measures related to the nature of Share interactions (specifically, their quantity, responsiveness, and reciprocity) we find that there is far more of this aspect of trust in the Hierarchy group than the Edge group.

We believe this contradiction is due to the powerful effect of the 17-wide posting (increased shared awareness) allowed in the Edge group that strongly limits the number of interactions (that aspect of trust) that is required to solve the problem. This kind of confusion highlights the need to propose hypotheses that deal with narrower aspects and/or definitions of trust as well as for us to do a better job isolating it from other factors; in this case from shared awareness.

Lastly, in order to make this research more accessible to others, it might helpful to add one or two explicitly hypotheses related to overall group performance (better or worse) related to organizational form (Edge or Hierarchy). The current work, of course, provides volumes of data dealing with this, but the existing hypotheses don’t make explicit predictions about performance related to organizational form.

Experiment Task

Our experience with live subjects shows that the difficulty of the task is approximately within range. While having subtasks with a range of difficulty is desirable, some of the more difficult subtasks could be correctly inferred without having full access to all the data. For experiment purposes, we may want to consider adjusting some of the factoids to improve scoring clarity See Appendix H on Factoid Sets for more information.

ICCRTS - ELICIT

Many more experiments need to be conducted to provide additional data. Results to date indicate the need for:

- More runs within a community
- More runs in differing communities
- Experiments to understand why some factoids are not shared or posted.

Analysis

The current experiment apparatus generates a significant amount of data. It would be possible to perform additional analysis of this data, particularly in the areas of shared awareness and trust.

When we have amassed more data sets it will be possible to do more formal statistical analysis of the data.

Possible Follow-on Experiments

Once the initial experiments are completed, there are many additional variables that could be explored. Follow-on experiments could:

- Introduce the use of reputation measures.
- Be based on new and/or more refined null hypotheses as discussed earlier.
- Vary the factoid distribution. For example, we could compare random distribution of factoids against distribution to those who have the most interest.
- Allow some of the factoids to contain incorrect or inconsistent information.
- Allow postings to be delayed, either for some individuals or for all subjects.
- Allow non factoid messages to be posted as freeform text. For example, we might encourage:
 - Leadership postings
 - Conclusion postings
 - Open-ended postings, although this may generate data that's difficult to analyze.

Additional Tuning of Initial Experiments

The following experiment conditions could be varied to provide additional data:

- Length of time between waves of factoid distribution.
- Determination of when the experiment trial is ended. Options include:

ICCRTS - ELICIT

- X moves after the last factoid is distributed.
 - X time after the last factoid is distributed.
 - When all subjects have identified a threat. This would provide more data points to understand the time required to accomplish the assigned task.
- Instructions to adjust motivation levels. Several variations of the instructions have already been generated.
- Source of the data to see the impact on distribution.

Appendix M: Web Site Support

This project has a companion Web site at <http://www.dodccrp.org/html3/elicit.html> that includes supporting materials for ELICIT. The Web site has three sections. Examples of material to be found in each area are listed below.

ELICIT Process Documents

- Videos for pre- and post-experiment briefings (Windows Media Player files)
- Moderator/Coordinator checklist
- Proctor Training agenda
- All supporting Factoid, Name and Country tables.

ELICIT Raw Data

- Summary of experiment trials.xls spreadsheet that catalogs the logfiles and survey files, etc.

ELICIT Analysis files

- Summary of experiment analysis files.xls that catalogs the sets of Excel workbooks that were used to process the raw log data to answer such questions as “Were Shares reciprocated?” And “Were all the factoids posted?”

Appendix N: Resources

Anklan, Patti, *Net Work: A Practical Guide to Creating and Sustaining Networks at Work and in the World*, (April 23, 2007).

Alberts, David S. and Hayes, Richard E.; *Power to the Edge: Command and Control in the Information Age* (May 2003).

Barkow, J., Cosmides, L. and Tooby, J.; *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford University Press (1992)

McCabe, Kevin, *Reciprocity and Social Order: What do experiments tell us about the failure of economic growth*" (2003).

McCabe, Kevin, Rigdon, Mary and Smith, Vernon; "*Mind, Cooperation and Social Exchange in Two Person Form Games*" (2002).