Defence Research and Development Canada
Recherche et développement pour la défense Canada

Copy No:

Joint Command Decision Support for the 21st Century Technology Demonstration Project

Lessons Learned from Experiment, Demonstration and Training Events

Richard Breton
DRDC Valcartier

Defence R&D Canada – Valcartier

Technical Report
DRDC Valcartier TR 2010-226
October 2011

Canada

Joint Command Decision Support for the 21st Century Technology Demonstration Project

Lessons Learned from Experiment, Demonstration and Training Events

Richard Breton
DRDC Valcartier

Defence R&D Canada – Valcartier

Technical Report
DRDC Valcartier TR 2010-226
October 2011

Principal Author

Original signed by Richard Breton

Richard Breton

Defence Scientist / C2 DSS

Approved by

Original signed by Patrick Maupin

Patrick Maupin

Section Head / C2 DSS

Approved for release by

Original signed by Christian Carrier

Christian Carrier

Chief Scientist

All events described in this report were sponsored by the JCDS 21 TDP under 15at.

In conducting the research described in this report, the investigators adhered to the policies and procedures set out in the Tri-Council Policy Statement: Ethical conduct for research involving humans, National Council on Ethics in Human Research, Ottawa, 1998 as issued jointly by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

Abstract

Over the course of the Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD) project, two exercises, one experiment and two demonstration sessions have been executed. This document presents the lessons learned from these five events. From these lessons learned, five guidelines have been identified. These guidelines should support the development of measurement protocols for military environments and appropriate metrics based on the constraints related to the environment in which the measurement should occur and the requirements related to each metric application.

Résumé

Durant le projet Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD), deux exercices, une expérience et deux sessions de démonstration ont été exécutés. L'objectif de ce document est de présenter les leçons apprises au cours de ces événements. À partir de ces leçons, cinq lignes directrices ont été définies afin de supporter le développement de protocoles de mesures pour les environnements militaires et l'identification de mesures appropriées basée sur les contraintes environnementales et les besoins particuliers de chaque mesure concernant son application.

This page intentionally left blank.

Executive summary

Joint Command Decision Support for the 21st Century Technology Demonstration Project: Lessons Learned from Experiment, Demonstration and Training Events

Richard Breton; DRDC Valcartier TR 2010-226; Defence R&D Canada – Valcartier; October 2011.

Background: Over the course of the Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD) project, two exercises, one experiment and two demonstration sessions have been executed. The objectives related to each type of events were stated as follow:

- Exercise: to train operators to execute Command and Control (C2) processes in joint, complex and time pressed environments;
- Experiment: to evaluate the quality of the support brought by potential technological solutions to increase process effectiveness;
- Demonstration: to show potential technological solutions to users.

The objectives of this paper are to identify lessons learned from these events, to define guidelines to support the development of measurement protocols appropriate for military environments and the identification of metrics specific for these environments.

Results: The lessons learned were related to the objective of the event, the type of metrics used, the environmental level of realism required and the availability of the participants. From these lessons learned, five guidelines were identified:

- Establishing the objective of the event (to train, to demonstrate, to study);
- Defining the to-be measured concept (task or process execution, or to understand the influence of an environmental or human factor);
- Establishing the importance of the measurement activities within the military event (primary versus secondary importance);
- Establishing the context in which the measurement activities take place (field trial, simulator or laboratory);
- Establishing the timing of the measurement activities (pre-event, during event or post-event).

These guidelines were also used in this document to evaluate potential metrics (observation, interview, questionnaire and objective measures) in terms of advantages/disadvantages or each measure, administration requirements and an evaluation of their applicability for each type of events (i.e., assessing which measures are best suited for a given type of event).

Significance: Measuring in military settings presents several challenges for the development of appropriate measurement protocols and selection of metrics. Often, measurement activities occur

in a large-scale event that has other primary objectives. Then, the measurement protocols must be fitted with the objectives, environmental constraints and opportunities of the event. In addition, the development of the protocols and the selection of metrics are often faced with important constraints in terms of participants' availability (limited number of participants and time to measure constrained).

While measurement is a challenge, the need for measuring performance is still growing with the importance and the complexity of technological support systems. It is critical to evaluate the impact of new systems in terms of human performance.

Thus, the identification of the lessons learned from the five activities performed within the JCDS 21 TD is surely an interesting step toward the development of appropriate measurement protocols and the selection of the best metrics suited for the situation.

Future plans: One major conclusion held in this report concerns the promising aspect of using objective measures in military settings. Objective measures present several benefits:

- technological setup not intrusive;
- data recording not obtrusive;
- direct performance measure (not biased by self-evaluation or evaluation from a tierce person);
- rapid to analysis;
- provide several measure.

However, they still require technological development. There is a need to develop an infrastructure that should support the data recording. Finally, a toolbox including metrics specifically developed for being used in military settings is required.

Sommaire

Joint Command Decision Support for the 21st Century Technology Demonstration Project: Lessons Learned from Experiment, Demonstration and Training Events

Richard Breton; DRDC Valcartier TR 2010-226; R & D pour la défense Canada – Valcartier; Octobre 2011.

Contexte: Durant le projet Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD), deux exercices, une expérience et deux sessions de démonstration ont eu lieu. L'objectif pour chacun de ces événements peut être défini comme suit:

- Exercice: Entraîner les opérateurs à effectuer des procédures de commandement et contrôle (C2) dans un environnement intégré, complexe et temporellement pressé;
- Expérience: Évaluer la qualité du support apporté par les solutions technologiques potentielles en matière d'amélioration de l'exécution des procédures;
- Démonstration: Présenter les solutions technologiques à des usagers potentiels.

Les objectifs de ce document sont d'identifier les leçons apprises de ces événements, de définir des lignes directrices afin de supporter le développement de protocoles de mesure ajusté aux environnements militaires et d'identifier les paramètres propres à ces environnements.

Résultats: Les leçons apprises concernent l'objectif de l'événement, le type de mesures utilisées, le degré de réalisme de l'environnement et le degré de disponibilité des participants. À partir de ces leçons apprises, cinq lignes directrices ont été énoncées:

- Établir l'objectif de l'événement (Entraîner, Démontrer, Étudier)
- Définir le concept à mesurer (Exécution d'un processus, d'une tâche, l'effet d'un facteur environnemental ou humain, etc.)
- Établir l'importance de l'activité de mesure dans l'événement militaire (importance primaire versus secondaire)
- Établir le cadre dans lequel la mesure sera prise (essais sur le terrain, dans un simulator, en laboratoire)
- Établir le moment de la prise de mesure (avant, durant, ou après l'événement)

L'identification de ces lignes directrices a contribué à l'évaluation de mesures potentielles (observation, interview, questionnaire et mesure objective) en fonction des avantages/inconvénients liés à chaque mesure, des besoins liés à leur administration et à leur capacité à être appliqués dans chaque type d'événement (quelle mesure est la plus appropriée pour ce genre d'événement.

Importance: La prise de mesures dans des environnements militaires pose des défis au développement de protocoles de mesure appropriés et la sélection de mesures. Souvent, les activités de mesure ont lieu dans des événements ayant des objectifs primaires plus important et

différents. Ainsi, l'établissement du protocole de mesure permet de les adapter aux objectifs de l'événement, les contraintes environnementales et les possibilités. En plus, le développement de ces protocoles est souvent confronté au problème de disponibilité des participants (en nombre et en temps alloué pour la mesure).

Tandis que la prise de mesure est un défi, le besoin de mesurer la performance est toujours grandissant avec l'importance et la complexité des systèmes de support technologique. Il est primordial de mesurer l'impact de nouveaux systèmes en matière de performance humaine.

Ainsi, l'identification des leçons apprises dans le cadre des cinq activités du projet JCDS 21 TD est sûrement un pas prometteur vers le développement de protocoles de mesure appropriés et la sélection des meilleures mesures possibles pour la situation.

Perspective: Une conclusion principale de ce rapport concerne l'aspect prometteur de l'utilisation de mesures objectives dans des environnements militaires. Ce type de mesure présente plusieurs bénéfices:

- l'environnement technologique n'est pas intrusif;
- le processus de la prise de données n'est pas obstruant;
- la mesure de la performance est directe (non biaisée par une auto-évaluation subjective ou de l'évaluation d'une tierce personne);
- les données sont rapidement analysées;
- fournit plusieurs types différents de mesures.

Cependant, cette avenue requiert encore des développements technologiques importants. Le besoin de développer une infrastructure pour la cueillette de données dans des environnements militaires est toujours présent. Finalement, le développement d'outils de mesure applicables spécifiquement aux environnements militaires est requis.

Table of contents

List of figures

List of tables

Acknowledgements

The author would like to thank Dr Adel Guitouni, Mrs Micheline Bélanger from DRDC Valcartier, Dr Sébastien Tremblay and Dr Robert Rousseau from Université Laval for their valuable comments and inputs and Dr David Bryant from DRDC Toronto for reviewing the manuscript.

1 JCDS 21 TD project

Over the course of the Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD) project, two exercises, one experiment and two demonstration sessions have been executed. The objectives related to each type of events were stated as follow:

- Exercise: to train operators to execute Command and Control (C2) processes in joint, complex and time pressed environments.
- Experiment: to evaluate the quality of the support brought by potential technological solutions to increase process effectiveness.
- Demonstration: to show potential technological solutions to users.

The next section provided an overview of the JCDS 21 TD project.

1.1 Project background

Defence Research & Development Canada (DRDC) conducted JCDS 21 TD project that sought to identify, refine, develop and integrate decision support and knowledge exploitation tools and demonstrate how these tool sets can significantly improve the ability of the Canadian Forces (CF) to respond to complex situations, such as those resulting from terrorist attacks and other asymmetric threats, in addition to other complex situations that call on the resources of the Canadian incident response environment.

More specifically, the objectives of JCDS 21 TD project were to:

- Understand the implications of net-centric operations within a Joint Inter-Agency Multi-National Public (JIMP) framework.
- Design and demonstrate a net-enabled collaborative environment that supports:
  - CF decision-making processes within a JIMP framework;
  - Exploitation of information and knowledge;
  - Collaborative working among distributed teams; and
  - Achievement of shared intent and decision superiority within a unified command framework.
- Develop operational and system requirements for related acquisition projects.
- Contribute to the Public Security Technical Program (PSTP) by sharing the results of studies and experimentation and collaborating on problems of common interest.

The project also sought to create a collaborative workspace to help CF experts work together effectively as well as coordinate with other government agencies, civil authorities and international allies when required. The collaborative workspace was designed to help experts integrate perspectives to better interpret the situation and the problem, identify candidate actions, formulate evaluation criteria, make decisions, and synchronize a diverse set of plans and actions.

In the JCDS 21 TD lifespan, five different events have been done (see Table 1). These events helped the JCDS 21 TD team to meet their objectives stated previously. However, as concurrent benefits, they also offered the opportunity to understand the complexity of running exercises, experiments or demonstration sessions in highly complex military environments such as the ones prevailing during the JCDS 21 TD events. Each of these events has been the subject of an experiment/exercise/demonstration report. The objective of this report is to identify lessons learned and draw conclusions regarding the execution of these events in complex CF environments.

Table 1 presents the five events included in the JCDS 21 TD project.

Table 1: JCDS 21 TD experiments, exercises and demonstration events.

Event	Date	Types	Relationship with JCDS 21 TD project
Ardent Sentry	May 2006	Training Exercise	Bi-National exercise (CAN-USA) in which JCDS 21 TD members contributed for the evaluation of C2 effectiveness of Canadian units.
Friendly Lance	June 2006	Training Exercise	The Friendly Lance exercise was part of a teaching program at the Canadian Forces College (CFC) Toronto that develops and practices the ability of the officers to execute the CFOPP. COPLanS, a tool to support the OPP and partially developed under the JCDS 21 TD project, was evaluated.
LiveSpace Familiarization sessions	May 2007	Demonstration	LiveSpace, an Autralian tool, has been offered to Canada for evaluation and collaboration under the TTCP between DSTO and DRDC. LiveSpace has been introduced to potential users of NDCC and JIIFC Det for evaluation. The original version of LiveSpace has been modified by the JCDS 21 TD project to fit with Canadian needs.
JCDS 21 EXP1	November 2007	Experiment	This exercise was run by CFEC in support of the Royal Canadian Mounted Police (RCMP) Integrated Security Unit (ISU) for the 2010 Olympics. A formal agreement with CFEC was established to allow JCDS 21 TD to validate during the exercise the C2I2 management processes of an operational joint headquarters in response to a major domestic event.
October Demo	October 2008	Demonstration	Formal JCDS 21 TD event in which new tools and concepts developed under the project help the operator in achieving better operational effectiveness.

These events can be subdivided into three distinct categories (experiment, exercise and demonstration). The following sections of this document provide a distinction between these three types of events.

1.2 Distinction between experiment, demonstration and exercise (training)

This section describes experiments, demonstrations, and exercises in general and highlights important differences among them.

1.2.1 Experiment

This section describes an experiment in terms of its objectives, the required level of realism of the testing environment, the capacity to use sophisticated and more intrusive types of metrics and the importance of the level of availability of subject-matter experts (SME).

1.2.1.1 Objective

The objective of an experiment is to establish cause-and-effect relationships between a specific manipulation and observed results in order to study a phenomenon in a more or less controlled experimental environment.

In 2006, The Technical Cooperation Program (TTCP) between Canada, Australia, United Kingdom and United States produced a Guide for Understanding and Implementing Defence Experimentation (GUIDEx) [1]. This GUIDEx proposed fourteen principles that should rule experimentations in defence environments. These principles are stated as follow:

1. Defence experiments are uniquely suited to investigate the cause-and-effect relationships underlying capability development.

2. Designing effective experiments requires an understanding of the logic of experimentation.

3. Defence experiments should be designed to meet the four validity requirements:

    a. Ability to employ the new capability;

    b. Ability to detect change;

    c. Ability to isolate the reason for change;

    d. Ability to relate results to actual operations.

4. Defence experiments should be integrated into a coherent campaign of activities to maximize their utility.

5. An iterative process of problem formulation, analysis and experimentation is critical to accumulate knowledge and validity within a campaign.

6. Campaigns should be designed to integrate all three scientific methods of knowledge generation (studies, observations and experiments).

7. Multiple methods are necessary within a campaign in order to accumulate validity across the four requirements.

8. Human variability in defence experimentation requires additional experiment design considerations.

9. Defence experiments conducted during collective training and operational test and evaluation require additional experiment design consideration.

10. Appropriate exploitation of modeling and simulation is critical to successful experimentation.

11. An effective experimentation control regime is essential to successful experimentation.

12. A successful experiment depends upon a comprehensive data analysis and collection plan.

13. Defence experiment design must consider relevant ethical, environmental, political, multinational, and security issues.

14. Frequent communication with stakeholders is critical to successful experimentation.

In summary, these fourteen principles are intended to cope with the complexity of military environment and the complexity of the measurement process.

1.2.1.2 Level of realism of the experimental environment

The level of realism of an experimental environment can vary over a relatively broad spectrum, from experiments done to study human information processing activities (basic research) to applied studies. Figure 1 (adapted from [2]) presents a spectrum of different types of experiments.

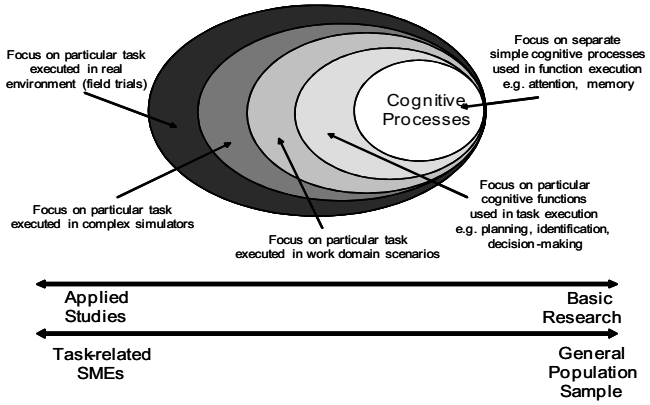

Figure 1: The Experimental Spectrum adapted from [2].

Basic research is illustrated on the right side of Figure 1. The objective of this kind of research is to gain knowledge on the human information processing activities. Consequently, it requires a very controlled experimental environment in which those activities are isolated. The task used is selected on the basis of its capacity to study a specific cognitive process. Since it is not related to a specific environment, the realism of the task can be compromised to the benefit of a better control over the variables of interest. Because the experimental task is simple and does not require specific expertise, a sample representing the general population is used.

Field trials are shown on the left side of Figure 1. In field trials, the focus of the experiment is not the study of a given cognitive process but rather the execution of a given task (i.e. decision-making in C2 environment). The goal of the experiment is to evaluate the human performance in the task execution. Consequently, it becomes critical that the experimental setup to study the task execution be as realistic as possible. To conduct field studies, cognitive and experimental psychologists are involved in the experiment as well as cognitive engineers and SME. SME are often used to develop and validate work domain scenarios. Field trials can be very costly, time and resource consuming and provide very little control over the variables of interest.

Basic research and field trials represent both ends of the experimental spectrum. Basic research focuses on the control of the variables of interest in the experimental setup at the expense of its realism. Field trials require high level of realism and then, may reduce in most occasions the control the experimenter has over the experiment. The recent technology improvement in computers has brought the development of microworlds that may be seen as an interesting compromise between basic researches and field trials. For instance, Jobidon, Rousseau & Breton [3] have used a microworld called “Save the Whale” to study the importance of temporal parameters in the selection of a control mode during a decision-making task execution. Such microworlds increase the level of realism; increase the number of parameters that can be considered in the manipulation and keep the level of control over the variables of interest at an appropriate level.

Experiments in military settings often involve the evaluation of new capacities with respect to their potential to support a given task. Consequently, such evaluation process requires a certain level of realism (ecological validity) without compromising the control over the variables of interest. This is the challenge experimenters are faced with in military settings but microworlds and high-level simulators are certainly an interesting avenue to help the experimenters to develop efficient real and controlled experimental setups.

1.2.1.3 Types of metrics

Metrics are used to gather data that allow the study of a given phenomenon (basic research) or a task performance (field trials). Several types of metrics are used to collect data:

- Interviews
- Questionnaires
- Observations
- Data recording (objective measures)

Interviews (structured, semi-structured or open) are used to probe participants about a specific subject. This type of metric is very flexible, easy to use and can lead to the collection of a large amount of information. In a structured interview, the practitioner uses a pre-defined set of questions. Semi-structured interviews are more flexible in the sense that only a portion of the questions are pre-determined. In those types of interview, the practitioner can direct the focus of the interview and can also use further questions that were not originally part of the planned structure. Open interviews allow the interviewer to explore, on an ad-hoc basis, different aspects of the subject. There is no pre-defined set of questions or structure. The more an interview is unstructured, the more it will rely on the interviewer's expertise about the subject of interest. For more information, see Stanton, Salmon, Walker, Baber and Jenkins [4].

Another method to collect information is through questionnaires. Questionnaires offer a flexible way to collect large amount of data from large sample. In comparison with interviews in which a one-to-one session is required, questionnaires can be administered simultaneously to almost unlimited numbers of people. Queries in the questionnaire can take several formats: multiple choice, rating scales, paired associates, ranking, open-ended questions, closed questions and filter questions (for more details, see [4]). Following the questionnaire's administration, the data analysis can be quick and straightforward particularly if the questionnaire was administered by computer. However, new questionnaires, by definition, lack a large data set that could be used in validation. Then, their reliability and the validity can be questionable.

Another method to collect data is through observation. Obviously, observations are only possible for observable (physical and verbal behaviours). Here are few examples of what can be observed:

- Individuals performing a given task (the steps included in the task execution);
- A technological system performing a task;
- The level of interaction between an operator and a system while performing a task;
- The communication between co-performers;
- The impact on environmental factors (time pressure, uncertainty) on the human and/or the system performance;
- The errors made by the human.

Drury [5] suggests five different types of information that can be elicited from observations:

- Sequence of activities (i.e. flowchart);
- Duration of activities;
- Their frequencies;
- Fraction of time spent in states; and
- Spatial movement.

Data from observations provides a "real-life" insight into the activity performed. However, it requires qualified observers (sometimes SME) that can capture the subtlety of a task execution. Also, observations are biased by the observer's personal experience and background. This problem can be solved by using several observers and correlating their observations. However, this increases the cost in terms of resources and also may increase the problem of intrusiveness

(presence of several observers in the experimental environment) related to observational data collection.

The emergence of microworlds and the development of more sophisticated technological environments offer new possibilities for data recording. Such technological environments offer the possibility to record eyes movement (with eye trackers), mouse-clicks, keystrokes, chat transcripts, etc. without necessarily requiring intrusive technological recording devices. From these data recordings, a large number of objective measures can be recorded and correlated. However, an understanding of the observed behaviour (for instance, mouse-clicking) in terms of the impact of this behaviour on the overall task performance is required. Such comprehension should be supported by theoretical background.

Based on the constraints related to military settings, questionnaires and observations have been traditionally the favoured way to collect data. Questionnaires are a quick and easy means to gather data. Observations require the availability of SME and the possibility to locate those SME in the experimental settings. However, advances in technology offered new possibilities to collect objective measures. The advantage of objective measures is that they are not based on participants' self-assessment and are unobtrusive. With the objective of assessing the role of cognition for the support of C2 tasks, Lafond et al [6] proposed objective metrics and experimental protocols. Results showed promising avenue to evaluate task performance in complex military environment. Their protocols offer great control over the variables of interest without compromising the level of realism of the experimental situations (unobtrusive measures).

1.2.1.4 Availability of participants

According to Kirk [7], specifying the number of subjects required for an experiment is often one of the more puzzling problems in experimental design. A representative sample size is one that provides an estimation of the overall population with a certain level of confidence (most of the time 95%). Such a level of confidence means that by doing the same experiment 100 times with a different population sample, the same observations from the data analysis could be made 95 times. Consequently, to represent the overall population, the number of participants must be sufficiently high. Obviously, the number of participants required will be affected by the number of conditions in the experiment. For example, let's say that we want to study the effect of time pressure on a monitoring task. We may establish, in our experimental protocol, three levels of time pressure (low-medium-high). To adequately compare the three conditions and to be able to extrapolate the findings to the overall population, an adequate number of homogenous participants need to be assigned to each condition. If the number of conditions is increased, the number of participants should be increased accordingly.

Problems related to the availability of participants become more important when the focus of the study is shifted to the execution of a given task (the left end of the experimental spectrum of Figure 1). To study a given task execution, several level of expertise (from novice to expert) may be necessary. Consequently, it reduces considerably the pool of potential participants. Two factors must be considered in the problem of the availability of the participants:

- The number of participants required in the experiment;
- The duration of the experiment.

As an example to illustrate the impact of these two factors (number and duration) let's consider that we want to evaluate the impact of a given support system on a given C2 task execution. Here is a list of requirements to develop an adequate experimental design:

- One group of experts is assigned to the condition (with the new tool);
- Another group of experts is assigned to the condition (without the new tool);
- Appropriate metrics to evaluate the task performance (time to execute, accuracy, etc).

In this type of experimental design, the homogeneity of both groups is critical. In order to rule out the possibility that any observed difference between the conditions is caused by a difference in the level of expertise between the groups (one group being more expert on average than the other), the level of expertise between the groups must properly balanced. The question of how experts differ from novices is a central concern for human engineering [8]. This concern raises the difficulty of defining precisely the notion of expertise. One may define expertise by the acquisition of specific knowledge that affects information processing and performance. Nevertheless, the difficulty to define precisely the expertise and the boundaries between the different levels makes the constitution of groups with comparable levels of expertise very difficult and arbitrary. One way to solve this problem would be to use this specific experimental design:

- Group 1 executes the task with the new tool and then re-execute the task without the DSS;
- Group 2 executes the task without the tool and then re-execute the task with the DSS.

In this within-subject design, the conditions with and without could be compared on the same set of experts. Also counterbalancing the moment at which the new tool is available in the experiment would rule out the possibility that the results being explainable by a training effect (for instance, if they start without the tool and re-execute after the same task with the tool). While this type of experimental protocol is very popular in basic research, it may be even more vulnerable to the problem of availability of participants from expert populations. It may be practically difficult and time consuming to gather enough experts for constituting two groups to play twice the same experiment.

Experiments in defence environments ask for the development of experiment protocols with clear objectives and appropriate metrics in order to maximize the availability (number and duration) of the experts. In fact, this problem is specifically addressed by the principles # 8 in the fourteen principles stated in the GUIDEx (Section 1.2.1.1).

1.2.2 Demonstration

This section describes demonstration sessions in terms of their related objectives, the required level of realism of the testing environment, the capacity to use sophisticated and more intrusive types of metrics and the importance of the level of availability of SME.

1.2.2.1 Objective

The objective of a demonstration is to showcase some new technologies that could be introduced in an operator environment to support the execution of his/her tasks. Breton, Paradis and Roy [9] stressed the importance of introducing technological testing in the development of a support system. A demonstration event in which technological systems are presented to potential users can also be referred to as "prototyping" in system engineering. Chapanis [10] describes prototyping as essentially a method of simulating the functions and behaviours of user-system interfaces with the capability of rapidly changing interfaces features. Similar to prototyping, demonstrations involve unfinished products that can be still modified based on users' feedback. Both demonstrations and prototyping can be used to gather feedback on these following aspects (for more details, see [10]):

- Verify or validate users' requirements and concepts;
- Evaluate alternative designs;
- Determine compliance with requirements for user or system performance;
- Identify problems of usability or functionality;
- Generate specifications of requirements for human-computer interfaces.

The benefits of prototyping or demonstration are that it reduces the number of design changes, brings the user in the loop in the design process, increases the likelihood of acceptance by the users, and provides an opportunity to get feedback from the users. Obviously, prototyping and demonstrations lay on the level of realism of the simulation. While both activities are similar in their execution, prototyping may intervene earlier in the design process. Demonstrations may be used only at the end of the design process to showcase the finished technological product to the targeted users.

1.2.2.2 Level of realism of the demonstration

Prototyping and demonstrations require fairly high levels of realism to fully meet their objectives. In Figure 1, the first three types of activities in the experimental spectrum from the left end of the figure should be favoured to execute prototyping or demonstrations. As mentioned previously, in these three types, the focus of interest of the experimental activities is shifted from the understanding of specific cognitive processes or functions to the understanding of a task execution. Generally speaking, the main objective of prototyping or demonstrations is to either evaluate the quality of a system to support a task execution (evaluate the product) or to demonstrate the capability of a given system to support a task execution (showcase the product).

These three types of experimental activities displayed in Figure 1 vary in terms of the complexity of the experimental environment. The middle layer of the figure represents situations where the focus of the interest is on the execution of a particular task in a given work domain scenario. The next one on the left focuses on particular task executed in more complex simulations. The last one at the very end of the figure represents situations where the focus of interest is on the execution of a given task during field trials. At the early stage of prototyping, the middle layer using a simple scenario should be favoured. The scenario should be developed to ensure that the functionalities that are the object of the prototyping activity are appropriately stimulated. With more complex

simulations, it is critical to establish clear relationships between specific functionalities in the prototype and observed results. With complex simulations, it could be difficult to attribute the observations of a given result with a specific functionality. This situation is even more important with field trials where multiple extraneous variables may contaminate the results. As a result, prototyping may be best suited for the middle layer of the experimental spectrum.

As mentioned previously, demonstration sessions could be used at the end of the design process to showcase specific functionalities or the overall capability of a system in the context of the execution of a given task. While prototyping may be best suited for the middle layer of the Figure 1, demonstrations are compatible with more complex and realistic experimental setup. In fact, to demonstrate the validity from a human performance or an operational perspective, one may use an environment that would simulate as much as possible the reality of the operational environment.

In summary, prototyping is used earlier in the design process and then, may require sufficient level of control in the experiment to link the observed results with specific functionalities included in the prototype. Demonstration requires higher level of realism to validate the appropriateness of a given system in the task execution.

1.2.2.3 Types of metrics

As stated previously, both demonstrations and prototyping can be used to gather feedback on these following aspects:

- Verify or validate users' requirements and concepts;
- Evaluate alternative designs;
- Determine compliance with requirements for user or system performance;
- Identify problems of usability or functionality;
- Generate specifications of requirements for human-computer interfaces.

The specific objective of a demonstration is to showcase some new technologies that could be introduced in an operator environment to support the execution of his/her tasks.

Obviously, the metrics used should be in line with the objectives of the prototyping or demonstration activities. Table 2 presents a summary of potential metrics that could be used to meet prototyping and demonstrations objectives.

Table 2: JCDS 21 TD experiments, exercises and demonstration events.

Prototyping	
Objective	Potential metrics
Verify or validate users' requirements and concepts	- Interviews: ask potential users what they would need to better execute a given task - Questionnaires : Administrate a questionnaire to gather information on what potential users would need to better execute a given task - Observations: Observe the potential users executing a given task in order to define potential needs and requirements
Evaluate alternative designs	- Interviews: ask potential users to provide their evaluation of different prototype alternatives - Questionnaires: Administrate a questionnaire that allow the ranking of different prototype alternative on specific aspects of the system - Observations: Observe different pattern of behaviours that would indicate differences between the prototype alternatives - Data recording: Record specific data such as time to execute a given task and the number of steps, mouse-clicking, keystrokes required to execute a given task with the different prototype alternative in order to compare these alternatives
Determine compliance with requirements for user or system performance	- Interviews: Ask potential users if the system meets their expectations - Questionnaires: Administrate a questionnaire to evaluate the quality level of the system in respect to the users' expectations - Observations: Observe specific behaviours that would suggest that the user of system performance is or is not optimal and relate that level of performance with the level of compliance with the requirements - Data recording: Record specific data such as time to execute a given task and the number of steps, mouse-clicking, keystrokes required to execute a given task that would suggest that the user of system performance is or is not optimal and relate that level of performance with the level of compliance with the requirements
Identify problems	- Interviews: Ask potential users if the system is user-friendly and if the

of usability or functionality	encountered usability or functionality problems - Questionnaires: Administrate questionnaire in order to evaluate the level of usability and functionality of the system - Observations: Observe potential problems in respect to the usability and functionality of the system - Data recording: Record specific (i.e. long period of time before getting information, too many steps, confusing functions, etc) data that would suggest usability and functionality problems
Generate specifications of requirements for human-computer interfaces	- Interviews: Ask potential users about their needs in terms of interface (i.e. number and type of screen, specific displays in information presented, type of interaction with the system (mouse, keyboard, voice, etc.) - Questionnaires: Administrate questionnaire in order to gather information on users' needs in respect to the interface of the system - Observations: Observe users executing the task with their actual systems or prototypes in order to identify potential problems with the interface
Demonstration	
Objective	Potential metrics
To showcase some new technologies that could be introduced in an operator environment to support the execution of his/her tasks	- Interviews: Ask users what they think, their feelings about the potential of the new technology - Questionnaires: Administrate questionnaire to evaluate the users perception about the new system - Observations: Observe specific behaviours that would suggest positive or negative users' feelings about the system - Data recording: Record data (timeliness, number of steps, accuracy, completeness, correctness, etc) that would backup conclusions about the quality level of the new technology

1.2.2.4 Availability of participants

Prototyping and demonstrations are different than formal experiments in which specific numbers of participants are required to reach a minimal level of statistical significance. Most of the time, these activities do not require statistical procedures. Generally speaking, the objective of both

prototyping and demonstrations is to gather feedback, feelings and information that would impact the design of a potential system. This general objective necessarily influences the number of participants required but, more importantly, the level of expertise of the participants.

To gather good information and feedbacks about the quality of the prototypes, the participants' level of expertise should be as comparable as the one of the potential users of the system. Preferably, the potential users of the system should be the one used in the prototyping sessions.

During demonstration sessions, the participants should be the potential users of the systems. These sessions should be seen as a sell-pitch in which the potential of the systems is displayed to the users.

Because of the importance of the level of expertise of the participants, one concern is the availability of the participants limiting the number of prototyping sessions (number of prototypes tested) and the length and the timing of the demonstrations sessions.

1.2.3 Exercise

This section describes exercises (training) in terms of their related objectives, the required level of realism of the testing environment, the capacity to use sophisticated and more intrusive types of metrics and the importance of the level of availability of SME.

1.2.3.1 Objective

The objective of an exercise is to train actual or future operators, decision-makers, or experts o optimally execute a given or a set of tasks. Consequently, during an exercise, the purpose is not to gather information to influence the design of a given support system or to showcase a given system. Instead, the purpose is to make the individuals comfortable in the execution of the task and the use of the available systems.

Nevertheless, with appropriate information gathering, observations or data recording processes, exercises are excellent opportunities to collect data to influence the development of training programs, as well as identify potentials problems from a system or user performance perspective and thus influence the design of new support systems.

From a Canadian Forces perspective, the objectives of exercise are:

- To train their members in the execution of a given or set of task;
- To train their members for the optimal use of available systems;
- To develop the expertise of their members.

From a scientific perspective, the exercise offers the opportunity to fulfill these objectives:

- To gather information on the development of training program;
- To understand how expertise is developed;
- To make the distinction between experts and novices;

- To identify potential problems in the task execution and the use of available systems;
- To identify design requirements for future systems;
- To identify team reorganizations requirements.

1.2.3.2 Level of realism of the exercise

The challenge for the scientific community is to meet those objectives without interfering with the exercise event. Alberts and Hayes [11] state that exercises greatly restrict behaviours and alternative approaches, thus preventing the ability of key variables to be controlled. One critical element of exercises is their very high level of reality. At the very left end of Figure 1, we describe field trials where the control over the variables of interest is compromised for the sake of reality of the trials. This compromise is even more severe in exercises. In the development of an exercise, the goal is to reproduce at the highest level of fidelity the same conditions that should prevail during real life situations. Consequently, intrusive data recording strategies that could compromise the fulfillment of this goal must be avoided.

1.2.3.3 Types of metrics

Alberts and Hayes [11] note that the results of most exercises are kept closely held to avoid embarrassing individuals who may have made errors as a part of learning and training process. In fact, the participants' feeling of being evaluated during the exercise could have a negative impact on the perceived level of reality of the exercise.

Nevertheless, given the value of exercises as opportunities to collect data, it is worthwhile to address this issue. One way is to sanitize the data by protecting individual's identity. Table 3presents a list of potential metrics with suggestions on how to make their use comply with the need for non-intrusive metrics.

Table 3: Metrics included in data collection plan (DCP).

Metrics	Application
Interviews	Interviews are possible, but only at the end of the exercise. Interviews could also be made at the end of each day if the exercise is run over many days. However, it is critical that the content of the interview does not suggest that the participants is evaluated or provide any clue about the exercise itself. The goal of interviews should be to collect information about participants' feelings, problems related to the task execution or system performance, and information required.
Questionnaires	Questionnaires can also be administrated at the end of the exercise or at the end of each day if they meet the same conditions (not suggesting any

	evaluation or providing any clues) as interviews. The goal of questionnaires should be to collect information about participants' feelings, problems related to the task execution or system performance, and information required.
Observations	Observations are possible during the exercise. However, to not disrupt the course of the exercise, observers should be as silent and non-intrusive as possible. It is important that observers, with their behaviours, do not provide clues to participants about what will happen in a near future. For instance, with the expectations of a reaction following a pre-determined event (that the observer knows), if observers are moving to observe a given position, their behaviours could suggest to this participant that he/she will have something to do in a near future. In addition, the number of observers should also be very low. Then, it may be a very challenging task for observers to collect good data specially if there are several participants in the exercises. This latter problem stresses the need for having SME as observers. SME should at least be able to identify positions of interest during the exercise. The goal of observation should be to identify behaviours suggesting problems related to the task execution or system performance, and information required.
Data Recording	The opportunity to use data recording devices is closely related to the nature of the exercise. In other words, the exercise environment must allow the use of data recording device. With computers, it is possible to record almost all interactions of participants with the support systems. However, this results in very large pool of data that takes very extensive time and resources to analyze. Also, it is critical that data recording process being not apparent to the participants. The goal of data recording should be to record any data suggesting problems related to the task execution or system performance, and information required.

1.2.3.4 Availability of participants

Because exercises are simulating real life operations, the number of participants should be based on the number of positions filled during real operations. Consequently, of the more participants required, the greater the potential difficulty problem of participants' availability. For instance, it may be very difficult to mobilize several experts at a specific period of time and for several days.

The expertise of the participants is also a critical factor. For an exercise held at the beginning of a training program, the expertise of the participants may be relatively low. However, when exercises are closed to the deployment of the participants in real life operations, the expertise level must be as high as possible.

2 Evaluation of the JCDS 21 TD events

The following table (Table 4) presents a summary of all JCDS 21 TD activities based on:

- Their objective of the event;
- The type of metrics used;
- The participants' availability and accessibility (capacity to interact with participants):
  - low: not accessible during and after the event
  - medium: not accessible during the event
  - high: always accessible before, during and after the event
- The level of realism of the simulation; and
- The level of intrusiveness allowed for the metrics application.

Table 4: Summary of the JCDS 21 TD events.

Activity	Objective	Type of metrics	Participant requirement (accessibility)	Realism	Level of intrusiveness
Ardent Sentry 2006	Training	- Observation - Questionnaires	Several participants for several days (low)	Very High	Very low
Friendly Lance	Training	- Observation	Several participants for several days (medium)	Very High	Very low
LiveSpaces	Demo	- Observation - Questionnaire - Data recording	Several participants for 2 days (high)	Medium	Medium
JCDS 21 TD EXP 1 (Pegasus Guardian I)	Performance evaluation	- Observation - Questionnaire - Data recording	Several participants for several days (medium)	High	Medium
October Demo	Demo	- Observation - Questionnaire - Data recording	Several participants for several days (high)	Medium	Medium

The following sections present the lessons learned from these five activities.

2.1 Ardent Sentry 06

2.1.1 Descriptive File

Type of event	Exercise
Objective	Training
Required level of realism of the event	Very high: It was critical that the level of realism of the exercise was higher as possible in order to replicate the operational conditions in which the trainees will have to execute their tasks in the operational environment. As a result, the capacity to control any variables of interest during the exercise was somehow challenged.
Type of metrics planned by the JCDS 21 team	- Observation - Questionnaire
Availability of the participants (number of participants and time period available)	For the exercise: several participants took part of the exercise over a period of 4 consecutive days. For the questionnaire: participants were mostly not available.

2.1.2 Background

The Ardent Sentry series of exercises was sponsored by the Chairman of the US Joint Chiefs of staff. These exercises were conducted by the North American Aerospace Defence – US Northern Command (NORAD-USNORTHCOM) and supported by the US Joint Forces Command. These exercises had a training purpose for NORAD-USNORTHCOM in homeland security and defence processes.

Canada participated previously in other Ardent Sentry exercises and was a partner in the development and the execution of Ardent Sentry 06 (AS06). The exercise objectives related to the Canadian participation were:

- Provide and coordinate incident management support by CF in support of Canadian civil authorities;
- Coordinate US incident management planning and response with Canadian counterparts;

- Conduct cross-border civil support operations with Canada; and
- Exercise Operation NOBLE EAGLE (ONE) procedures in Canadian NORAD region.

In the Ardent Sentry 06 exercise, the JCDS 21 TD team has been given the mandate of evaluating C2 effectiveness of Canadian units, particularly Canada Command. In addition to this objective, the team took the opportunity to validate their measurement tools to be used during future experiments or exercises such as PEGASUS GUARDIAN.

2.1.3 Data Collection Plan

This section provides a summary of the data collection plan used by the JCDS 21 TD team during AS 06. Based on the conditions prevalent in the AS06 exercise, the measurement tools needed to be as less intrusive as possible and they needed to rely on observation activities and post-hoc questionnaires. These measurement activities were planned in AS06:

- Two Pre-Exercises surveys (See Annexes A and B);
- The Command Team Effectiveness (CTEF) administered before and after the experiment (See Annex C);
- Two surveys (before and after the exercise) evaluating the tools used, shared situational awareness and commander's intent (Annexes D and E).

Daily observations were also planned regarding the communication between operators. Observers were also tasked to collect some information such as:

- Physical Layout of Ops Centre
- Operational Tempo
- Team Structure
- Decision Making and Collaboration

2.1.3.1 Assigned locations within the JCDS 21 team members

Ardent Sentry was played at different locations. Table 5 identifies the locations and recommending JCDS 21 staff for each.

Table 5: JCDS 21 TD experiments, exercises and demonstration events.

Locations	Number of assigned peoples for low level measures	Number of assigned peoples for low level measures
Canada Com (Ottawa)	2	4
JTFA (Halifax)	2	0
JTFC (Toronto)	2	0
GOC (Ottawa)	1	0
SJS (Ottawa)	1	2

It was planned that the "low level measures" observers were dedicated to the site as shown and were responsible for the completion of the observation plan and the conduct of interviews. Other observers divided their time between the three locations in Ottawa during the week assisting the other observers and interfacing with J7 or senior officers involved in the Exercise.

2.1.3.2 Observer responsibilities

- Before the exercise, observers were getting familiarized with exercise materials, Concept of Operations (CONOPs), Standard Operating Procedures (SOP), plans and procedures and the exercise DCP toolkit.
- Each observer received a copy of the DCP toolkit that contains data collection packages to guide observations.
- Observers were expected to attend the Observer Orientation Briefing (prior to the beginning of each day) that included a detailed review of exercise activities and the scenario. This briefing was a key time for observers to ask questions and ensure that they completely understood their roles and responsibilities.
- Observers were given any updates from exercise planners on changes to plans and procedures.
- During the exercise, the observers' primary duty was to observe and record what the players did. After the exercise, that information was used to determine whether the expected training outcomes were achieved and to identify strengths and opportunities for improvement.
- Observers did not simply attach themselves to the team, section, or organization leader. The best place to be was wherever the observer could see and hear the action.
- An observer did not take a position where he or she might be a distraction or interfere with the exercise play.

2.1.3.3 Observers record keeping

Observers were asked to keep an accurate written record of what they saw and heard. To be reliable, they had to take notes as players took actions and made decisions.

Notes were planned to contain the following:

- Who (by name or position) performed the action or made the decision;
- What occurred (the observed action);
- Where (the location) the action or decision took place;
- When (the time) the action took place;
- Why the action took place or decision was made (the trigger);
- How they performed the action or made the decision (the process).

2.1.3.4 What to look for and record

Because numerous events occurred simultaneously, it was expected to be difficult for the observers to record all the action. To make recording the action manageable, observers were made aware of what the important events were, and to eliminate superfluous information, they were provided the kind of data most useful for exercise evaluation. Records of important events were expected to include the following (note that this information can be captured with the communication matrix):

- Message in: An individual or group receives information from somebody outside of their physical location. Messages can be sent via radio, telephone, e-mail, fax or another means other than face-to-face conversation. If known, indicate if the message is in the Main Scenario Event List (MSEL).
- Message out: An individual sends information to another individual or group of people outside of their physical location. Messages can be sent via radio, telephone, e-mail, fax or another means other than face-to-face conversation.
- Discussion: A conversation involving several people.
- Decision: An individual or group arrives at a conclusion or makes a specific determination. Decisions might be made following a discussion or can be made independently. Decisions often, but not always, lead to directives.
- Directive: An individual gives an order or specific direction to one or more people. An individual in a position of authority often, but not always, gives a directive. Directives can be given in person or via other types of communication such as radio or telephone.
- Movement: An individual, group or piece of equipment relocates.
- Activity: An individual or group performs a specific, clearly definable action or function.
- Inject: Information, including directives, instructions, and decisions that are provided by exercise controllers to exercise players. Injects can be written, oral, or televised and can be transmitted via any means (e.g., fax, phone, e-mail, voice, radio, or sign).

Other things to record were expected to include the following:

- Initiating scenario events (including when players first detect abnormal conditions);
- Deviations from plans and implementation procedures;
- Timeliness of critical actions;
- Monitoring and assessing scenario events;
- Command and control at the scene;
- Creative player problem solving beyond current plans and implementation procedures;
- Plans or procedures that affect player efforts;

- Equipment issues that affect player efforts.

2.1.4 Lessons learned in AS06

The objective of the AS06 activity was to train future operators to:

- Provide and coordinate incident management support by CF in support of Canadian civil authorities;
- Coordinate US incident management planning and response with Canadian counterparts;
- Conduct cross-border civil support operations with Canada; and
- Exercise ONE procedures in Canadian NORAD region.

In this exercise, the JCDS 21 TD team had two objectives:

1. Evaluating C2 effectiveness of Canadian units, particularly Canada Command; and

2. Validating their measurements tools to be used during future experiments or exercises such as PEGASUS GUARDIAN.

2.1.4.1 Lessons learned related to metrics

Observation: Metrics used in AS06 were questionnaires and observations. The observation process was challenged by a limited number of observers allowed in the room. The evaluation of C2 effectiveness of Canadian units has been very difficult for the JCDS 21 TD team for several reasons:

- Because of the difficulty to conduct interviews with the participants (availability and accessibility problems), most data were collected from subjective observations. In order to draw validated conclusions from subjective observations, a given phenomena should have been observed several times by different observers in different conditions. Such conditions would have required several observers and repeated exposures to same situations. Unfortunately, during AS06, these conditions were not possible.
- Because very few observers were allowed in the room, it was very difficult for the observers to cover all the action. It was even more difficult for them to record all relevant actions for all positions in the team.
- Observers were not experts. Consequently, it was difficult for them to capture all the sense in the action and to find elements of interest during the scenario play.

Recommendation:

- Clearly identify a set of pre-defined behaviours to be recorded. This categorizing process would be done in collaboration with SME.
- Train the observers to recognize this set of pre-defined behaviours.
- Negotiate with stakeholders in order to increase the number of observers.

Observation: The JCDS 21 TD team prepared for that exercise several questionnaires in order to test their robustness and validity. Unfortunately, they had very few opportunities to validate them for future administrations for the next JCDS 21 TD events. However, based on what they observed during the exercise, they conclude that:

- The number of questionnaires should be reduced;
- The length of the questionnaires should be reduced;
- Because of the time constraints and the availability of participants, it may be very difficult to administer the same version of a given questionnaire to the same participant. Consequently, repeated measures to compare different key dimensions (level of shared SA, common intent, etc) at different moments during the exercise may not be feasible.

Recommendation:

- Negotiate with stakeholders the planning of a formal administration period (for instance, at the end of each day).
- Reduce the length and number of questionnaire by improving the content of the questions.

2.1.4.2 Lessons learned related to the realism of the exercise

Observation: AS06 required a very high level of realism in order to recreate the operational environment in which the operators have to execute their tasks. The evaluation of C2 effectiveness of Canadian units has been very difficult for the JCDS 21 TD team for some reasons related to the level of realism of the experiment:

- Participants were not available to conduct open interviews at the end of the simulation. Although they received an instruction from the commander to feel free to participate in these interviews, they were not required to do so. Without a formal assignation, the participation level has been very low.
- Also, the high level of realism constrained considerably the number of observers allowed in the exercise environment.

Recommendation:

The level of realism is a constraint that needs to be taken into consideration during the development of the DCP. Facing with such constraint, here is a recommendation to reduce the impact of this constraint and to maximize the gain:

- Utilize SME as observers. SME could rapidly detect what behaviours need to be recorded. That would reduce considerably the number of observations made that may not be relevant.

2.1.4.3 Lessons learned related to the availability of the participants

Observation: There were several participants involved in AS06. In addition, their participation was required for several consecutive days. However, little room was left for the JCDS 21 TD team to interact with these participants.

Recommendation:

- Identify and focus on key positions pre-determined in collaboration with the stakeholders.
- Negotiate with stakeholders for formal interview sessions.

2.2 Friendly Lance Exercise

2.2.1 Descriptive File

Type of event	Exercise
Objective	Training
Required level of realism of the event	Very high: It was critical that the level of realism of the exercise was higher as possible in order to replicate the operational conditions in which the trainees will have to execute their tasks in the operational environment. While to experimental groups (one using COPLanS and the other not) were formed, the JCDS 21 TD team had very limited control over the course of the exercise.
Type of metrics planned by the JCDS 21 team	- Observation - Questionnaire
Availability of the participants (number of participants and time period available)	For the exercise: several participants took part of the exercise over a period of 4 consecutive days. For the questionnaire: participants were mostly not available.

2.2.2 Background

In a Command and Control environment, the success of planning activities allows the decision-maker to be one step in front of the opponent. Successful planning may overcome a lack of resources and manning. Based on the importance of planning activities in tactical and operational environments, military organizations devotes considerable time and efforts to develop approaches

to allow them to execute the planning process in an optimal way. The Canadian Forces Operational Planning Process (CFOPP) is a systematic approach to analyzing a situation, bringing staff expertise to bear on the relevant factors, narrowing courses of action, obtaining the commander's approval, and developing the detailed annexes necessary to produce an executable plan. It is the planning process used by the CF at the strategic and operational level.

DRDC Valcartier has developed a computer-based system intended to support the CFOPP [12]. This system, called COPLanS, has been designed to provide the ability to plan an operation in a net-enabled environment using integrated collaborative tools. COPLanS is an integrated, flexible suite of planning, decision-aid, and workflow management tools aimed at supporting a distributed team involved in the military operational planning process. It is a client-server as well as a Web-based application. The system offers functions to design, manage and synchronize multiple concurrent battle rhythms at the strategic and, operational levels and to a limited extent at the tactical level. It helps synchronize the staff workflow, documents automatically the decision-making process and allows the replaying of the decision-path. The planning tools help the staff to sketch courses of action (COA) on maps, to perform time and space synchronization, to manage resources and capabilities, to manage ORBAT and to perform logistics analyses. The decision-aid tools help the staff to rationalize the process, to improve the COA evaluation and comparison, and to rapidly produce documents to support the Commander's decisions. COPLanS offers multi-level collaborative tools including Chat, White Board and On Map Planner. A context sensitive search engine is available to browse past similar operations and recall plans and lessons learned from the database.

In order to evaluate COPLanS as a support tool for the CFOPP, it was decided to collaborate with CFC Toronto, which is the Canadian excellence center for teaching the CFOPP to the military officers. Different exercises are conducted as part of their teaching program to develop the ability of the officers to execute the CFOPP. Friendly Lance, which is one of these exercises based on the Atlantis scenario, was identified as being appropriate to provide the context needed to evaluate COPLanS.

JCDS 21 TD members took part to this exercise with the mandate of:

- Demonstrating and evaluating COPLanS as a collaborative planning system at the operational level;
- Identifying requirements related to the achievement of Decision Superiority in the planning of operations while being in a Joint, Net-Enable Collaborative Environment.

2.2.3 Data Collection Plan

This section provides a summary of the DCP used by the JCDS 21 TD team during the Friendly Lance exercise.

The Friendly Lance exercise provided an opportunity to empirically evaluate COPLanS following an experimental approach (experiment), an opportunity to demonstrate the capabilities of COPLanS to potential users (demonstration) and to collect their feedbacks for future development (prototyping). Thus, the JCDS 21 TD team leveraged this exercise for experimenting and prototyping and demonstrating with COPLanS.

To meet their objectives, the JCDS 21 TD team identified three different groups defined as follow:

- The first one, hereafter called the potential user group (JOPG5), was composed of military officers (n=5) mostly from CANADACOM, CEFCOM and CANOSCOM.
- The second group, the COPLanS group (JOPG4), was constituted of Canadian Force College (CFC) Toronto students (n=17) executing the task using COPLanS.
- The control group (JOPG3) was composed of CFC Toronto students (n=16) that were ask to execute the task without using COPLanS.

The evaluation of the impact of COPLanS on the CFOPP execution involved only JOPG3 and JOPG4. Both groups were constituted similarly. They have been tasked to execute all phases of the CFOPP with the only difference that one group executed the task as it is actually performed in real operations and the other with the support of COPLanS.

The first objective was to demonstrate and evaluate COPLanS as a potential system to support the collaborative planning activities at the operational level. To achieve this, different hypotheses were verified during this exercise:

Hypothesis 1: COPLanS maintains staff synchronisation while executing the different Operational Planning Process (OPP) activities thereby improving the tempo of the decision-making process.

Hypothesis 2: COPLanS maintains near-real time staff updating on new information (information sharing) and associated deductions thereby improving the decision-making process.

Hypothesis 3: COPLanS, with its collaborative environment, should help performing collaborative mission analysis in a structured and rational manner thereby improving the quality of the staff assessment and the information brief.

Hypothesis 4a: The use of COPLanS should have an impact on the quality of these tasks execution. It should result in improving the COA quality.

Hypothesis 4b: It should contribute to rationalizing the decision-making process resulting in the decision quality improvement.

The DCP in the evaluation of the impact of COPLanS faced two important constraints. First, the OPP tasks were executed in three separate rooms and there were several sources of extraneous variables. Second, the limited availability of participants prevented the experimenters from doing repeated measures during multiple trials. Moreover, the primary goal of this exercise was to evaluate their performance following the course. Then, to avoid any interruption of their work, it was not possible to conduct interviews with these groups. To overcome these constraints, as much as possible, a strategy involving quantitative and qualitative measures was used.

The data collection strategy gathered three different methods to record information for the evaluation of the hypotheses. These methods are listed as follow:

- Post-hoc questionnaire to evaluate the level of confidence of officers about the developed CFOPP outputs. It also provided a subjective evaluation of the impact of COPLanS on the execution of the CFOPP stages (Annex F).

- Targeted observations during the exercise made by an observation team.
- Gathering of exercise artefacts (i.e. exercise injects, commander's directives, etc) and produced outputs (i.e. Mission Analysis Brief, Information Brief, Decision Brief).

JOPG5 was constituted of Subject-Matter Experts. The objective with this group was to gather information about the potential of COPLanS as a support system for the CFOPP. The potential user group was used to identify requirements related to the achievement of Decision Superiority in the planning of operations.

The following grid indicates who conducted the previous strategies. DND Learn, the CFC Toronto portal for exercises was used to host the questionnaires. The feedback gathering tool of COPLanS was used to record comments from the two groups using COPLanS. COPLanS team gathered artefacts and JCDS 21/OR team collected observations and conducted interviews.

Table 6: Data collection strategy for each group

Strategy\ Group	Questionnaire	Observations	Artefact Gathering	Interviews	COPLanS feedback gathering
Potential User Group	DND Learn	JCDS 21	COPLanS team	JCDS 21/OR	COPLanS
COPLanS Group	DND Learn	OR	COPLanS team	None	COPLanS
Control Group	DND Learn	OR	COPLanS team	None	None

2.2.3.1 Observation Team

In the data gathering process, the JCDS 21 team and Operational Research (OR) scientists were working jointly to collect information for evaluating the effectiveness of COPLanS in the support of the CFOPP. Data were also gathered for identifying new requirements for the achievement of Decision Superiority in the planning of operations in a Joint, Net-Enable Collaborative Environment. Here was the constitution of the JCDS 21 and OR team:

- JCDS 21 team: Richard Breton, Renee Chow, David Smith, Kevin Trinh
- OR team: David Connell, Shaye Friesen

2.2.4 Hypotheses validation strategy

For each hypothesis identified previously, an assessment strategy is described as follow:

Hypothesis 1: COPLanS maintains staff synchronisation while executing the different OPP activities thereby improving the tempo of the decision-making process.

Measurement strategy:

- Observations were used to determine the level of **staff synchronisation** during the execution different OPP activities. Synchronization is defined by the timing in information requests and information transmission.
- Observations were used to determine the time required to produce the different outputs. Furthermore, since DND learn is the medium to exchange information for the exercise, the time the outputs are posted on the DND learn was recorded.
- Questionnaires were used to determine if COPLanS had a positive impact on the **tempo of the decision making process**.

Hypothesis 2: COPLanS maintains near-real time staff updating on new information (information sharing) and associated deductions thereby improving the decision-making process.

Measurement strategy:

- The observation team recorded the **latency detection time on new information** (as perceived by the staff officer).
- Questionnaires were used to determine if COPLanS had a positive impact on t**imely information sharin**g.

Hypothesis 3: COPLanS, with its collaborative environment, should help performing collaborative mission analysis in a structured and rational manner thereby improving the quality of the staff assessment and the information brief.

Measurement strategy:

- The observation team performed a post hoc qualitative comparison of the Mission Analysis Brief produced by the COPLanS group with the one produced by the control group to identify the difference in the level of **staff assessment** between these two groups.
- Questionnaires were administered to determine if COPLanS had a positive impact on the **staff assessment**.

Hypothesis 4a: The use of COPLanS should have an impact on the quality of these tasks execution. It should result in improving the COA quality.

Measurement strategy:

- The observation team performed a post hoc qualitative comparison of the Information Briefs/Decision Briefs produced by the COPLanS group with the one produced by the control to determine the difference in the level of **quality of the COA** between these two groups.
- Questionnaires were administered to determine if COPLanS had a positive impact on the **quality of the COA**.

Hypothesis 4b: It should have contributed to rationalizing the decision-making process resulting in the decision quality improvement.

Measurement strategy:

- The observation team performed a post hoc qualitative comparison of the Decision Brief produced by the COPLanS group with the Decision Brief produced by the control group to determine the level of **quality of the decision** between these two groups.
- Questionnaires were administered to determine if COPLanS had a positive impact on the **decision quality**.

2.2.5 Lessons learned in Friendly Lance

This JCDS 21 TD activity was particular in a sense that during a CF exercise (Friendly Lance), the JCDS 21 TD was given the opportunity to showcase one of its application (COPLanS) and to evaluate its impact within an experiment. The objectives for the JCDS 21 TD team were:

- Demonstrating and evaluating COPLanS as an OPP collaborative planning system at the operational level.
- Identifying requirements related to the achievement of Decision Superiority in the planning of operations while being in a Joint, Net-Enable Collaborative Environment.

The Friendly Lance activity was a blend between an experimentation in which a system is evaluated based on its capacity to support the human performance to a given task and a demonstration session in which a prototype is shown in order to gather feedbacks from SME for future development.

2.2.5.1 Lessons learned related to the metrics

The GUIDEx provides some principles to support the evaluation of the metrics used to cover the evaluation of COPLanS as a support system.

Principle: Defence experiments are uniquely suited to investigate the cause-and-effect relationships underlying capability development.

Observation: Clearly, the Friendly Lance exercise was not suited to uniquely investigate the cause-effect relationships underlying capability development. Instead, the JCDS 21 TD took the opportunity to leverage this activity to evaluate one of its applications. Unfortunately, there were problems:

- While it was possible to define two comparable groups (one using COPLanS and the other not), it was not possible to repeat the comparison process under different conditions at different moments. Then, results are only suggesting that at this specific moment in the time, a group of persons performed similarly or differently than the other. It is not possible to reach statistically validated conclusions from this isolated evaluation in the time.

- The JCDS 21 TD team had a very limited control over the length and the quality of the training period. In fact, it has been suggested that the quality of the training was not necessarily similar between both groups.

Recommendation: The Friendly Lance exercise provided the opportunity to regroup the participants in two distinct groups, one using COPLanS and the other not. However, considerations should be made on the assignation of the participants to the different groups in order to make sure that both groups are similar. More control of the training period should be necessary.

Principle: An iterative process of problem formulation, analysis and experimentation is critical to accumulate knowledge and validity within a campaign.

Observation: The experiment portion of the Friendly Lance exercise offered a one-shot opportunity to the JCDS 21 TD team to test and evaluate the COPLanS application. Consequently, an iterative process with problem formulation, pre-testing, analysis and testing was not possible. This limited considerably the testing and evaluating capacity of the experiment. It is for that specific reason that the results of the experiment can only be used to identify general tendencies.

Recommendation: For that specific reason, it would be essential to provide more scientific background to COPLanS development in order to better support the hypotheses stated in the experiment.

Principle: Designing effective experiments requires an understanding of the logic of experimentation;

Observation: The JCDS 21 TD team included several Human Factor experts specialized in the development of experimental protocols and the application of metrics to evaluate the human performance. Unfortunately, the team had to fit their experimental strategy in order to disrupt as less as possible the course of the exercise. Consequently, they had very few opportunities to administrate their data collection tools. The inclusion of a valid but more intrusive and controlled experimental protocol was not simply possible in this exercise.

Principle: Defence experiments should be designed to meet the four validity requirements;

- Ability to employ the new capability
- Ability to detect change
- Ability to isolate the reason for change
- Ability to relate results to actual operations

Observation: With the use of COPLanS, the first requirement was met. However, because of the aforementioned reasons, the experimental protocol did not allow clearly to detect changes in the effect. To do so, it would have been necessary to include several conditions (for instance, the presence of events that make to situation more or less complex at different periods of time) and to compare the performance of both groups (with and without COPLanS) as a function of these different conditions. This would have lead to the establishment of relationships between the

observed performance of both groups when using, and when not using, COPLanS. Such an ideal experimental protocol would have allowed relating the change to the postulated cause (stated in the hypotheses). Finally, statistically valid results reflecting changes in the effect could have been applicable to the operational forces in actual military operations. Clearly, in this experimental protocol, the problem preventing the meeting of these four validity requirements was a problem of low statistical power (not enough data collected).

Principle: Defence experiments conducted during collective training and operational test and evaluation require additional experiment design consideration;

Observation: This is exactly what happened with the evaluation of COPLanS during the Friendly Lance exercise. This principle states the importance of ensuring a certain level of experimental validity in order to make sure that the observed results can be extrapolated to operational situations. In the evaluation of COPLanS, while two experimental groups were formed, the application of appropriate metrics was difficult.

Recommendation: The problem with questionnaires and observations is related to their subjective nature. It would be necessary to use more objective measure such as logs, response time, eye trackers, etc. These measures are not intrusive and are seen as more reliable than questionnaires and observations.

Observation: Unfortunately, the type of subjective metrics used (questionnaire) is highly affected by several factors such as participants' mood, state of readiness, motivation, question understanding, etc that could influence the results. While these factors could be counterbalanced among both groups, with only one administration during the data collection process, there is still potential for data contamination. Also, it may be possible that the sensitivity of the metrics (its capacity to identify any change or difference between groups) may be not sufficient. Consequently, it can be very difficult, with such limited measures to draw conclusions.

Recommendation:

- It would have been necessary to administrate the questionnaire in many occasions in order to identify trends in the results.
- It would have been necessary to verify the psychometrics properties of the questionnaire which are:
  - Reliability (how stable or consistent a measure is);
  - Validity (how well it actually measures what is claims to measure);
  - Sensitivity (the ability of a measure to identify any changes due to the impact of different factors (such as the use or not of COPLanS to execute the task).

Observation: The Friendly Lance exercise offered also the opportunity, to the JCDS 21 TD team, to collect feedback about the COPLanS application from a group of experts potentially users of COPLanS (JOPG 5).

This part of the exercise can be seen as a demonstration or prototyping in which feedback about the following aspects are collected:

- Verify or validate users' requirements and concepts
- Evaluate alternative designs
- Determine compliance with requirements for user or system performance
- Identify problems of usability or functionality
- Generate specifications of requirements for human-computer interfaces

This activity allowed the JCDS 21 TD team collect interesting feedbacks on the different functionalities of COPLanS specifically for the identification of potential problems of usability and functionality. However, only 5 experts from three different organizations have been interviewed. Consequently, while interesting, we may not be able to generalize their comments to all experts from their respective organizations.

Recommendation: More experts from these three organizations would have been required to collect statistically valid information. These feedbacks can only be considered as the opinion of these 5 different experts and only tendencies can be identified.

2.2.5.2 Lessons learned related to the level of realism

The GUIDEx provides some principles to support the evaluation of the level of realism of the experiment that evaluates COPLanS as a support system.

Principle: Human variability in defence experimentation requires additional experiment design considerations;

Observation: Results showed that both groups perceived similarly their level of expertise and their level of knowledge about executing the OPP. However, results also suggested that the group using COPLanS may not have had sufficient training before using the application. Consequently, results could be explained by the lack of training for one group. It would have been better to provide enough training sessions for the COPLanS group.

Principle: An effective experimentation control regime is essential to successful experimentation;

Observation: This principle refers to the classical reality/control trade-off in the development of experimental protocol. As stated previously in this document, very controlled experiments may suffer from poor reality level. Conversely, experiments executed in real environments may lack precise control of experimental variables. Advanced simulators and microworlds may be good compromises with adequate level of reality and good control over the variables of interest. In this exercise, the JCDS 21 TD team did not have any control over the course of the exercise. They had to make their measurement strategy totally compliant with the exercise.

Recommendation: In the Friendly Lance exercise, the level of realism of the exercise was relatively high. Problems concerned more the control aspect of the experiment. In order to preserve the level of realism, limited control was given to the JCDS 21 TD team. As mentioned previously, the use of objective measures could have prevented this problem by keeping the level

of realism of the experiment as required but provided to the JCDS 21 TD team more control over the collection of data.

2.2.5.3 Lessons learned related to the participants' availability

Observation: Several participants took part to the Friendly Lance exercise. Concerning the experiment part of the JCDS 21 TD activities (comparison between two groups), there were enough participants in both group to establish statistically valid comparisons. However, repeated questionnaire administration would have been beneficial.

The JOPG5 group (SME) was available throughout the exercise. However, the five SME came from different organizations with different needs expectations and backgrounds.

Recommendation:

- Negotiate with stakeholders more questionnaire administration periods.
- Get more SME from the different organizations.
- Identify parallels and distinctions between the SME organizations in order to provide a sense to the collected feedbacks.

2.3 LiveSpaces familiarization sessions

2.3.1 Descriptive File

Type of event	Demonstration
Objective	To present an application and its functionalities
Required level of realism of the event	Medium: Relatively high because the participants executed their real duty (meetings) using the application. However, the JCDS 21 TD team had the capacity to interact with the users during their tasks execution.
Type of metrics planned by the JCDS 21 team	- Observation - Questionnaire - System performance metrics (Data recording)
Availability of the participants (number of participants and time period available)	For the exercise: several participants took part of the exercise over a period of 3 consecutive days. For the questionnaire: participants were mostly available.

2.3.2 Background

LiveSpaces, an Autralian product from Defence Science & Technology Organization (DSTO), has been offered to Canada for evaluation and collaboration under The Technical Collaboration Program (TTCP) between DSTO and DRDC. In the JCDS 21 project, a "Buy and Try" activity was sponsored to define specifications for a collaborative working capability aimed at the Operational Command and Fusion Centers.

Hence, in the context of this Buy and Try activity, three familiarization sessions with potential users were organized at Joint Information and Intelligence Fusion Capability (JIIFC) (May 14-16, 2007). The main objective of these familiarization sessions was to demonstrate the LiveSpaces environment and its capabilities to potential users. Three groups were asked to conduct their meeting (no simulation) using the LiveSpaces environment. Participants were mostly from National Department Command and Control (NDCC) and JIIFC Detachment (JIIFC Det) organizations. Electronic questionnaires were administrated at the end of each session. Observers from the Fujitsu Company (2) and DRDC-Valcartier (1) were gathering notes and observations during these familiarization sessions.

2.3.3 Data Collection Plan

This section provides a summary of the DCP used in the LiveSpaces familiarization sessions.

The three familiarization sessions were not considered experiments in which variables were controlled and measured. Instead, they were seen as opportunities to gather participants' first impressions about the tools and their potential to support collaborative working.

The familiarization sessions were done with participants from JIIFC Det and NDCC organizations. In the first session, members of Command Network Working Group (Comd-Net WG) were asked to brainstorm on the Command View requirements and planned evolution for the next two years. The objective was to identify and prioritize future enhancements and future opportunities. The brainstorming activity was mainly supported using the TeamThink capability of LiveSpaces. SharePoint was also used throughout the session for recording action items, notes and participants and making available relevant documentations to everyone.

In the second meeting, participants from JIIFC Det, NDCC, Canada COM, CEFCOM and CANOSCOM were asked to review their current and post annual posting season (2007) under the requirement of the new Common Command Support Element (CCSE) structure in function of the their capability to support the command requirements of the transformed CF. While participants from NDCC cancelled because of schedule conflicts, the task was executed as expected following a brainstorming approach using the TeamThink application. SharePoint was also used throughout the session for recording action items, notes and participants and making available to everyone relevant documentations.

The JIIFC Project is responsible for developing and delivering to the CF an information fusion capability that will be resident on the CF SECRET Network. During the afternoon session, a team had the objective of articulating the high-level requirements for this architecture. They used a brainstorming approach with TeamThink. They also used other LiveSpaces applications such as SharePoint, ScreenSharing and E-Beam.

In literature [13-14], there are multiple factors that are listed as potential variables affecting collaborative working. The most recurring and important ones are 1) communication; 2) information distribution; and 3) coordination. In other words, to efficiently collaborate, team members must communicate, the information must be properly distributed to the team members requiring this specific information (at the appropriate moment in an appropriate format) and their actions must be coordinated. While not exhaustive, these three factors seem to be part of the most critical ones for collaborative working. Questionnaires were constructed to assess the impact of LiveSpaces on these three dimensions.

In the questionnaires, questions related to communication focused on:

- the exchange of task-related information;
- communication of feelings, opinions and thoughts (not necessarily task-related);
- the exchange of feed-backs.

Participants were asked to rate on a "1-to-7" scale (1= rarely; 7= all the time) what they **usually** experience executing this type of task **without** LiveSpaces and what they **actually** experience **with** the LiveSpaces environment supporting the execution of a same type of task.

In the questionnaires, questions related to information distribution activities considered these following criteria on the rating scale:

- timeliness of the information (1= not on time; 7= on time);
- clarity of the information (1= not clear; 7= very clear);
- correctness of the information (1= not correct; 7= totally correct);
- and completeness of the information (1= incomplete; 7= totally complete).

Participants were asked to rate on these scales what they **usually** experience executing this type of task **without** LiveSpaces and what they **actually** experience **with** the LiveSpaces environment supporting the execution of a same type of task.

In the questionnaires, questions related to coordination activities considered these following criteria on the rating scale (1= rarely; 7= most of the time):

- know who is doing what;
- others know what you are doing;
- wait for other inputs;
- others wait for your inputs;
- jobs are complementary;
- jobs are supportive;
- providing feedback.

Participants were asked to rate on these scales what they **usually** experience executing this type of task **without** LiveSpaces and what they **actually** experience **with** the LiveSpaces environment supporting the execution of a same type of task.

2.3.4 Lessons learned in LiveSpaces familiarization sessions

The main objective of these familiarization sessions was to demonstrate the LiveSpaces environment and its capabilities to potential users.

2.3.4.1 Lessons learned from the metrics

The LiveSpaces familiarization sessions were considered to be demonstration or prototyping sessions. The objective was to demonstrate to potential users the capabilities of a given application and to collect valuable feedbacks for future development. To fulfil that need, the JCDS 21 TD team used a data collection strategy including observations and questionnaire.

Observation: The JCDS 21 TD team had another objective with the use of the questionnaire. The objective was to validate the content of the questionnaire to evaluate the quality of system to support the human performance for future use during the next JCDS 21 TD activities. In that sense, the LiveSpaces familiarization sessions provided a great opportunity to validate the questions based on these dimensions:

- number of questions included in the questionnaires
- clarity of the questions
- their validity

The administration of the questionnaire did not lead to statistically valid results as too few participants responded to the questionnaire. However, the JCDS 21 TD team were able to collect valuable feedback concerning the quality of the questionnaire. The length of the questionnaire and the clarity of the questions were favourably judged by the participants.

The questionnaire covered three important dimensions:

- Information distribution
- Communication activities
- Coordination activities

Again, the questionnaire used seemed to cover particularly well these three critical aspects in teamwork.

As mentioned previously, the primary objective of these sessions was to demonstrate the capabilities of LiveSpaces and to evaluate its capacity to support teamwork. The observations made by the observers provided very interesting results. Post-hoc analyses allowed the evaluation of LiveSpaces based on some design principles held from MacMillan et al [15; 16]. Other analyses were made based on the capacity of LiveSpaces to support leadership and adaptability behaviours within the team.

Recommendations: LiveSpaces, from a demonstration standpoint, was a very successful activity. Obviously, the participation of more SME and more time would have beneficial.

2.3.4.2 Lessons learned from the level of realism

Observation: In the context of a demonstration session, the level of realism was adequate. In fact, the participants were asked to execute their real life tasks with LiveSpaces. It was not a simulation. However, because they were not in a tactical situation, there were still able to interact with the observers.

2.3.4.3 Lessons learned from the participants' availability

Observation: In the context of the demonstration session, enough participants took part to the activity. In the context of the validation of the questionnaire, further participants would have been necessary.

2.4 JCDS 21 TD Experiment 1 (in conjunction with Pegasus Guardian Ex)

2.4.1 Descriptive File

Type of event	Experiment
Objective	To evaluate the team performance
Required level of realism of the event	High: The scenario used was realistic and the participants executed their tasks in a real operational environment (mobile lab). It was not possible for the JCDS 21 TD team to interact with the participants during the live play.
Type of metrics planned by the JCDS 21 team	- Observation - Questionnaire - System performance metrics (Data recording)
Availability of the participants (number of participants and time period available)	For the exercise: several participants took part of the exercise over a period of 4 consecutive days. For the questionnaire: participants were all requested to fill up the questionnaires.

2.4.2 Background

The aim of the JCDS 21 TD is to demonstrate a *Joint Net-enabled, Collaborative Environment* to achieve *Decision Superiority*. After the CF transformation and the orientation on Canada First,

the project was directed to focus on domestic operations (DomOps). The tenure of the Olympics in Vancouver in 2010 (V2010) became an excellent context to consider for training of DomOps for the CF.

This experiment was the first official testing and evaluation event held by the JCDS 21 TD team. The team had backup roles during Ardent Sentry 06 and Friendly Lance. LiveSpace was seen as an opportunity to demonstrate the capabilities of a system to potential users. However, it was not part of the original plan in the TD program.

The intent of this JCDS 21 TD event was to validate the Command and Control and Intelligence and Information (C2I2) management processes of an operational joint headquarter in response to a major domestic event. In particular, reactive C2I2 processes in response to calls for assistance were examined. Therefore, this first event aimed at assessing CF processes, procedures and C2 applications to support time sensitive decision making when dealing with DomOps.

A formal agreement with the Canadian Force Experimentation Centre (CFEC) was established in order to organise the JCDS 21 TD first event in conjunction with CFEC exercise Pegasus Guardian (EX PG). This exercise was run in support of the Royal Canadian Mounted Police (RCMP) Integrated Security Unit (ISU) for the 2010 Olympics. This event allowed JCDS 21 TD team members to validate their initial knowledge management and decision support concepts and technologies with respect to organisational and individual factors, situational awareness (SA) and operations planning and execution.

An Olympic 2010 scenario was used to simulate a complex decision-making situation in which decision makers are overloaded with information. This experiment simulated a deployed Joint Task Force Game Head Quarter (JTFG HQ), supported by the Joint Task Force Pacific / Games (JTFP/G) personnel. EX PG was supported by an unclassified stand-alone Local Area Network (LAN), including the JCDS 21 TD Test Bed, set up for the exercise by JCDS 21 TD staff.

This event allowed the examination of shared SA and information sharing issues within a CF HQ among team members supporting the command's decision cycle. The quality of decisions were assessed against a set of metrics developed by the JCDS 21 TD team and based on research findings in literatures from academia institutions, DRDC, TTCP and the North Atlantic Treaty Organization (NATO).

The JCDS 21 TD processes focused on the passage of information from the ISU and a joint command centre to create SA, to manage current operations, to respond to Request for Assistance (RFA) from the ISU, to respond to a Transfer of Authority (ToA) and its consequences, and to facilitate the assessment and planning of military responses beyond current events.

The experiment offered an opportunity for the collection of operational data to meet the aims of JCDS 21 TD, concurrent with an operational sequence to initiate the staff procedures of JTFG. JCDS 21 experiment 1 (JCDS 21 EXP1) focussed its analysis on the flow and management of information internal to the JTFG with regard to maintaining a shared SA, and responding to a RFA or a TOA.

The initial JCDS 21 TD and JTFG objectives for this first event were:

a. JCDS 21 TD to review and validate CF time-sensitive decision-making processes in response to requests from Other Government Departments (OGD): RFAs and ToAs;

b. JCDS 21 TD to assess shared SA, and respond to a RFA or a ToA. The assessment of the processes and the gap analyses were performed against a set of metrics developed by JCDS 21 TD Team and detailed later in this report;

c. JTFG to determine, on an initial basis, the processes and procedures that were required of JTFG as a functional tactical HQ, including an opportunity to identify desired changes to the organization;

d. JTFG to develop a better understanding of the linkages, lines of communication and flow of information through the JTFG but in particular between the Integrated Command Center (ICC) and JTFG;

e. JTFG to develop a better understanding of what portions of JTFG HQ need to be in the ICC; and

f. DRDC Command, Control, Computers, Communications, Intelligence, Surveillance and Reconnaissance (C4ISR) Mobile Lab was deployed to the West Coast for the EX PG. This lab allowed emulating existing DND baseline as well as demonstrating new concepts like knowledge discovery, geo-intelligence and knowledge packaging tools to assess their value-added to help achieving SA. It allowed JTFG to develop a better understanding of the DRDC mobile command trailer, its potential and its capability as well as demonstrating an initial idea of an integrated Command decision support environment.

JCDS 21 EXP1 was an unclassified event. All information used in developing the scenario and incidents has been taken from unclassified sources. However, some of the incidents represent possible operational scenarios so some of the storylines are considered sensitive and that portion of the information is protected accordingly.

2.4.3 Data Collection Plan

In the JCDS 21 EXP1, the experimental setup needed to recreate a realistic and complex enough environment to stimulate the JTFG team members. In addition, there was a need for creating an adequate simulated environment that allowed the collection of data.

In the context of the micro-world defined for the JCDS 21 EXP1, the components, defined as follow are providing the organizational structure:

- Reduced JTFG HQ: In this micro-world, a reduced deployed JTFG HQ (JTFG) plays its role as the main DND face for the game and interfaces with the ISU. Given the constraints of EX PG, all interactions between JTFG HQ and ISU are managed and simulated by the white cell. On exceptions, DND LO 2 communicates information to JTFG using C2 applications or other communications means.

- White Cell: The White Cell simulated all other potential players and organisations. It also simulated HI and LOW CON. A retired military flag officer played the role of the Commander.
- EX Control: The EX Control cell was essential to orchestrating all the events and the exercises events. It also filtered information exchange and made sure that no contamination occurred between CFEC Experiment and JCDS 21 EXP1. EX Control also provided oversight data collection and analysis activities.

In addition to the development of the micro-world detailed in the previous section, the JCDS 21 EXP1 included a set of experimentation methods:

- CPX: Command Post Exercise (CPX) consisted of creating realistic situations that engaged JTFG HQ. Events, triggers and background information were injected by the white cell in order to stimulate JTFG HQ to start appropriate C2I2 processes. The events and triggers were modulated in order to exercise different aspect of the processes identified and to collect data for analysis and simulation.
- Table-Tops: Table top discussions (i.e. "Table-Tops") were structured around the validation and assessment of existing C2I2 processes and ways for improvement (gaps and requirements). Table Tops followed CPX and reviewed play, draw conclusions and refined the models.
- Constructive simulation: C2I2 process validation and analysis were based on constructive simulation. "Constructive simulation is often referred to as "war gaming" since it bears some resemblance to table-top war games in which players command armies of soldiers and equipment that move around a board" [http://en.wikipedia.org/wiki/Simulation]. The C2I2 models were injected into simulation environment in order to further analyse gaps, perform "what-if" analyses and provide inputs for the table tops discussions.
- Questionnaires: a set data collection questionnaires were used to collect data for further analyses and validation of the C2I2 processes.

There were some constraints and limitations with this experimental setup:

- Injects production needed to be coordinated with the information sent to the team by the JIIFC ICC DND LO. This raised the importance of timing in the experiment. Injects needed to be introduced at very specific moment in the scenario.
- To keep the level of realism as high as possible, the data collection strategy needed to be as less intrusive as possible. The measurement strategy needed to be limited to observations and the application of questionnaires at the end of the day.

Metrics were based on the constraints and opportunities provided by the experimental setup, the participants' characteristics and the tools used.

To avoid an excessive level of intrusiveness of the metrics, the data collection strategy needed to rely on observations made by data collectors during the scenario play and questionnaires administered at the end of the day. Electronic recording of time spent on each application, electronic communications (chat and email transcript), etc was also possible.

2.4.3.1 Analysis of Situation Awareness

One mandate of the JCDS 21 TD team was to assess the quality of participants' SA and their understanding of the situation. Based on the opportunities and constraints defining the experimental environment, the measurement strategy was limited to the administration of questionnaires. These questionnaires, the Situation Awareness Rating Technique (SART), the Mission Awareness Rating Scale (MARS) and the Quantitative Assessment of Situation Awareness (QUASA) are presented respectively in Annexes G, H and I.

2.4.3.2 Analysis of Collaboration within the team

Based on the opportunities and constraints characterizing the experimental environment, these measurement activities were selected:

- **Observations:** Observations were made on the frequency of verbal communications between participants.
- **Electronic recording**: All electronic information exchange activities (emails and chats) were recorded. Note that these sources of data are not analyzed in this document. They could be the object of another set of data analysis in the future.
- **Questionnaires:** Questionnaires on the quality of the communication & collaboration among team members were administrated (see Annex I) to the participants at the end of each day.

In the JTFG cell located in the Mobile Lab, there were many face-to-face communications between the different positions recorded by two observers using a communication grid.

- From the content of those grids, it is possible to determine the frequency of communication between each position and type of information exchanged between each position.

Then, the content of the verbal communications was categorized following these categories:

- General Communications: It includes, for example, request for information, clarification, and request to take action, response statement, planning statement, decision statement, and factual statement.
- SA: This category is decomposed into statements related to the perception of the situation, understanding and projection of course of evolution of events.
- C2 processes: This category is limited to communications referring to activities related to RFA and ToA because C2 processes were not observed directly.
- Background knowledge: This category represents all communications related to background discussions about doctrine, background expertise, etc.

The recording activity covered the communications made from electronic devices (phone or radio) or via a network (chat, email and TITAN). More specifically, it included log of emails of each participant, log of chat sessions, and log of all VoIP communications between participants. Note that the analysis of these data is not presented in this document.

To cover the communication and collaboration aspect, a questionnaire was administered to all participants (see Annex I) at the end of each day. This questionnaire covered these topics:

- **The quality of the information exchange activities.** More specifically, it provided participants' subjective assessments for the timeliness of the information, the clarity of the information, the correctness of the information, and the completeness of the information.
- **The quality of the coordination activities among team members.** More specifically, it provided participants' subjective assessments about their capacity to know who is doing what in the task execution, to make other team members know what they are doing in the task execution, about the time they wait for others inputs, and about the time they make other team members wait for their inputs.
- **The quality of the collaboration among team members.** More specifically, it provided participants' subjective assessments about the level of complementary of each team member contributions, the level of support between each team member contributions, and the quality level of feedback.

2.4.3.3 Analysis of C2 applications

Finally, in this JCDS 21 TD event, C2 applications used by the participants were also evaluated from the administration of a questionnaire (see Annex J).

Following a question related to their actual level of experience with the three C2 applications (Command View, Mission View, Information Management System), each participant rated these applications of 5 different dimensions:

- Clarity: Question relating to the clarity provided by each of the three C2 applications. The rating utilized a five-point Likert scale format – 1: Strongly Disagree to 5: Strongly Agree.
- Timeliness: Question relating to the timeliness of information provided by each of the three C2 applications. The rating utilized a five-point Likert scale format – 1: Strongly Disagree to 5: Strongly Agree.
- Correctness: Question relating to the accuracy of information provided by each of the three C2 applications. The rating utilized a five-point Likert scale format – 1: Strongly Disagree to 5: Strongly Agree.
- Completeness: Question relating to the completeness of information provided by the three C2 applications. The rating utilized a five-point Likert scale format – 1: Strongly Disagree to 5: Strongly Agree.
- User Satisfaction: While the questions above provide useful information of their own, we created a new index of "User Satisfaction" for each of the three C2 applications. The index was based on these factors: accuracy, reliability, timeliness, relevancy, content, format, and ease of use. The User Satisfaction index combines the ratings on questions relating to clarity, correctness and completeness on the C2 applications questionnaire into a single measure. User Satisfaction is the average value given by each participant to these questions for each of the three C2 applications.

Each participant was asked to fill out the questionnaire at the end of the experiment (Day 3).

2.4.3.4 Data collection team

The data collection team was composed of five observers and one supervisor. All these data collectors are DND employees.

Data Collection Supervisor

- Tasks Assignment:
  - Prior to the experiment, the supervisor prepared all data collection materials (questionnaires, communication grids, Data Collection protocols);
  - Prior to the experiment, the supervisor briefed all participants about the data collection strategy;
  - During the experiment, he supervised all data collection activities;
  - During the experiment, he prepared/adjusted the questionnaires based on any changes occurring during the day;
  - At the end of each day, he participated to all debriefing sessions;
  - At the end of the experiment, he participated to the hotwash sessions with military personals;
  - After the experiment, he analysed the data collected during the experiment and prepared the documents.
- Location: in the White Cell and occasionally in the Mobile Lab

Two communication flows data collectors embedded within the HQ:

- Tasks Assignment:
  - Prior to the experiment, they participated to a basic training session on the C2 application (Monday morning). This gave them the opportunity to get familiar with the environmental context and the C2 applications;
  - Prior to the experiment, they assisted to the training sessions provided to the JTFG participants on the C2 applications (Monday afternoon). This gave them the opportunity to identification the role for each participant and to observe the group dynamic;
  - During the experiment, they recorded of all communication activities between each JTFG positions;
  - At the end of each day, they produced a daily report;
  - At the end of each day, they participated to all debriefing sessions;
  - After the experiment, they supported the data analysis by categorizing all communications activities recorded during the experiment into specific classes as reported in this document.
- Location: in the Mobile Lab

Two C2I2 processes observers

- Tasks Assignment:
  - Prior to the experiment, they participated to a basic training session on the C2 application (Monday morning). This gave them the opportunity to get familiar with the environmental context and with the C2 applications;
  - Prior to the experiment, they assisted to the training sessions provided to the JTFG participants on the C2 applications (Monday afternoon). This gave them the opportunity to identification the role for each participant and to observe the group dynamic;
  - During the experiment, they recorded all processes used to support the C2I2 activities. They observed the task execution via a camera located in the Mobile Lab. A camera transmitted the visual and auditory feeds from the Mobile Lab to the observers' laptop. Headset were used to hear properly the verbal exchanges between the participants;
  - At the end of each day, they produced a daily report;
  - At the end of each day, they participated to all debriefing sessions;
  - After the experiment, they supported the data analysis by providing more clarification on the processes observed.
- Location: in the White Cell

D/ORM DISB, Questionnaires manager

- Tasks Assignment:
  - Prior to the experiment, he participated to a basic training session on the C2 application (Monday morning). This gave him the opportunity to get familiar with the environmental context and with the C2 applications;
  - Prior to the experiment, he assisted to the training sessions provided to the JTFG participants on the C2 applications (Monday afternoon). This gave him the opportunity to identification the role for each participant and to observe the group dynamic;
  - During the experiment, he followed the scenario in order to verify which QUASA probes are relevant and should be presented at the end of the day;
  - At the end of each day, he managed the questionnaires administration;
  - At the end of each day, he participated to all debriefing sessions.
- Location: in the White Cell during the experiment and in the Mobile Lab during the questionnaires administration.

2.4.4 Lessons learned from the JCDS 21 TD Experiment 1

The JCDS 21 EXP1 was the first official testing and experimentation event held under the TD. Consequently, more resources and efforts were assigned to this event. This resulted in an impressive pool of data recorded and analysed. The next section provides a list of observations in

regards to the metrics used to collect the data, the level of realism of the experiment and the availability of the participants. Following this overview of the experiment, we identify lessons learned and make some recommendations.

2.4.4.1 Lessons Learned related to the metrics

The initial JCDS 21 TD and JTFG objectives for this first event were to assess shared SA, and respond to a RFA or a ToA. This section includes recommendations from observations made during the experiment and the analysis of some of the fourteen principles included in the GUIDEX.

2.4.4.2 Observations

Observation: In the DCP, there was no direct measure used for evaluating shared SA. Obviously, there is a strong relationship between sharing SA and exchanging information. Two questions included in the Comm/Coll questionnaire concerned the exchange of information activities within the team:

1. other team members knew who was doing what within the team:

2. other team members knew what I was doing during the task execution:

While these two questions cannot be considered as a direct measure of sharing SA (sharing SA cannot be simplified to exchanging information), it still can provide an assessment of the level of shared SA within the team. Results indicated fluctuations in the level of information exchange throughout the experiment that could be correlated with the level of comfort of the participants within the environment and the intensity of the crisis.

Recommendation: To adequately measure the level of shared SA within the team, there is a need to clearly define the concept of Shared SA. It could be difficult to develop a measure that encompasses all the activities executed when SA is shared among participants. Consequently, the measurement of sharing SA could be better done by identifying key functions (i.e. exchanging information) and identifying metrics that measure specifically that function (i.e. analysing the emails and chats exchanges among the team members).

Observation: Most metrics included in the DCP evaluating the level of SA of participants have been insensitive. Only QUASA presented some significant differences in the results throughout the experiment.

Recommendation: A potential explanation for the insensitivity of the SART and MARS metrics may lie in the fact that these metrics were covering too long periods of time (one day). Then, questions included in these questionnaires were referring to situations too general to allow insight into participants' SA. To support this hypothesis, Rousseau, Tremblay, Banbury, Breton, & Guitouni [17] suggest that subjective self-assessment metrics such as SART and MARS are mostly determined by generic declarative knowledge built from past experience and background. Participants would not necessarily refer to what occurred during the exercise to determine their ratings. MARS and SART measures of SA did not seem to be sensitive to changes and evolution

of the exercise. The difference with QUASA was that the questions were referring to specific events. Optimal use of QUASA would require probes to be inserted throughout the exercise rather than at the end of each day. While it would not be possible to freeze the experiment in order to administrate the QUASA probes, a good compromise would be to get QUASA probes in the middle of the day (at the lunch time) that would cover what happened during the morning and other probes at the end of the day for the afternoon activities. It would have been more beneficial to administrate these relatively short questionnaires more often during the experiment.

Observation: The experiment was managed with the MORAE environment. This type of environment allows the recording of several sources of information. For instance, to record precisely the detection of new events in the scenario, several data sources could be correlated. First, it would require analyzing all mails or chatting transcripts for all participants to observe records of new detections. Second, it would require the analysis of all keystrokes or mouse clicking following the occurrence of a new event in the scenario. Third, it would require the analysis of all face-to-face communications between team members.

Recommendation: While this type of correlation analysis was not done, it would be interesting to develop an analytical framework to provide a theoretical background to this analysis. Also, from the analysis of the communication content (face-to-face communications, mails, chatting transcripts), it could be possible to verify the quality of the interpretation of the participants as well as their capacity to predict the future status of the situation.

Observation: The social communication network provided an analysis of the communication activities within the team. However, the networks produced by this analysis provided information on the communication activities regardless of the situation (type of tasks) or content (type of content). It would still be possible to analyze the timeliness and the exactness of the communication activities through the mails and chatting transcripts. However, this type of analysis could be made only on a portion of all communications activities since most were made face-to-face.

Recommendation: All communication should be classified with respect to the related task or communication content. Applying a Social Network Analysis to the communication data was very informative about the interaction pattern of the JTFG. In fact, this type of analysis could be used to define the most appropriate or optimal physical setup of JTFG.

Observation: In addition, because of the restricted number of observers in the mobile lab, it was not possible to follow all communications that occurred between all participants and it was not always possible to record the type of content exchanged. Consequently, the analyses were performed on a non exhaustive set of data. Then, the results may potentially be biased by the observers' capacity to keep track of the communication, their ability to detect interesting communications flows and their understanding of the tasks.

Recommendation: Audio and visual feeds of the mobile lab should be recorded for post-hoc analysis of the communications by SME.

2.4.4.3 Principles

This event was considerate as an experiment. Consequently, some of the 14 principles included in the GUIDEx produced from the TTCP effort can be used to provide further evaluation of the event.

Principle: Defence experiments are uniquely suited to investigate the cause-and-effect relationships underlying capability development;

The objective of this experiment was exactly to investigate new processes, SOP or potential technological solution. However, it was not *uniquely* what occurred during the experiment. Some VIP tours were held within the experimental environment that could have challenged the quality of the experiment.

Recommendation: The term 'uniquely' is critical in this principle. As it can be seen in this document, most military experiments are executed in conjunction with other training or demonstration events. Consequently, it may be difficult to control the course of the experiment (control over the variable of interest) and to reduce the impact of extraneous variable such as the presence of VIPs during live play. It is important to try to follow this statement in order to adequately evaluate cause-and-effect relationships.

Principle: Designing effective experiments requires an understanding of the logic of experimentation;

For this event, hypotheses have been stated based on theoretical background. A complex and challenging scenario was developed. Enough participants were assigned to test these hypotheses. Unfortunately, the JCDS 21 TD team had very limited opportunities to collect observations (not enough observers allowed in the Mobile Lab). Also, the administration of the questionnaires was also constrained.

Recommendation: It would be beneficial to demonstrate the importance of the data collected from questionnaires to stakeholders. They need to understand what can be extracted from these sources of data.

Principle: Defence experiments should be integrated into a coherent campaign of activities to maximize their utility;

The JCDS 21 EXP1 leveraged greatly from the previous events. From these events, appropriate questionnaires were built and selected to maximize the data collection activities.

Principle: Frequent communication with stakeholders is critical to successful experimentation;

The JCDS 21 TD team established a very good relationship with the stakeholders by presenting them the importance and the benefits related to this event.

2.4.4.4 Lessons Learned related to the realism of the experiment

Observation: The experiment was also used as a training event for the participant somehow reducing its level of realism. In several occasions, the validity of the experimental setup has been challenged:

- Participants were liberally exchanging on SOP, task execution, etc.
- Participants were talking together during the questionnaires administration.
- A VIP tour occurred during the experiment. Then, visitors were going in and out into the mobile lab.
- The participants did not have formal training prior to this task assignment. Consequently, they were learning while executing the tasks. This could explain the difference in results between Day 1 and Day 2.

Recommendation: A better control of participants' communications would still be critical. Also, while VIP tours are important, it would be necessary to identify appropriate moments in the experiment for these events.

Observation: Because the metrics required being less intrusive as possible, the data collection strategy needed to rely on observations made by data collectors during the scenario play and questionnaires administered at the end of the day. These following restrictions limited considerably the possibility of data collection:

- There were only two data collectors (recording the face-to-face communication) allowed in the mobile lab. Considering the importance and the potential of the Social Network Analysis framework, it would be important, for future experiment, to increase the coverage of this type of communication.
- The questionnaires were only administered at the end of the day. Then, at the time of the administration, the events that occurred during the morning could have hardly been recovered. The questionnaires should be administered in the middle and at the end of each day.

Recommendation: These following recommendations could reduce the impact of these limitations:

- Negotiate with stake-holders for more on-site data collectors
- Record the visual audio feeds of the mobile lab for post-hoc analyses
- Negotiate with stake-holders for more questionnaire administrations

2.4.4.5 Lessons Learned related to the availability of the participants

Observation: The number of participants in the JTFG cell was limited to ten. While this number was sufficient for the Comm/Coll, C2 apps and QUASA questionnaires, it seemed not sufficient for the SART and MARS.

However, there was a more important problem concerning the number of participants taking part of the experiment. Usually, in any experimental protocol, participants need to be considered as being equivalent. Statistical procedures are used to counterpart any discrepancies in the participants (age, sex, culture, etc). In our experiment, participants were not equivalent since they

were assigned to different roles and positions. Based on these positions, these following aspects were varied:

- Level of centrality within the team: some positions were expected to be more central than others;
- Sociometric status: some positions were expected to be busier than others;
- Use of C2 applications: the need for specific C2 applications was not the same for all positions;
- Information distribution: access to all information was not necessarily required for all positions.

Consequently, while we were pooling the results of all participants together, the positions filled by these participants could not being judged as being equivalent.

Recommendation: To reduce the impact of these major distinctions between the positions, an adequate experimental setup would have required many different JTFGs in order gather the results of each position altogether. For instance, let's say that we have 10 participants and five of them have access to a specific application and the five other don't. Then, by comparing these two groups (with and without the application), we could have better assess the impact of the application on this specific application. Unfortunately, the limited availability of military peoples makes such requirements very difficult to meet.

2.5 October Demo

2.5.1 Descriptive File

Type of event	Demonstration
Objective	To showcase JCDS 21 TD applications
Required level of realism of the event	Medium: The scenario was almost the same as the one used during the PG / Bronze experiment with some minor modification. While the Mobile Lab was still used, other participants were positioned in a tent. During the live play, several VIPs were admitted for a tour. There was a lot of distraction, specifically for the participants in the tent. It was possible for the JCDS 21 TD team to interact with the participants during the live play.
Type of metrics planned by the JCDS 21 team	- Observation - Questionnaire - System performance metrics

Availability of the participants (number of participants and time period available)	For the exercise: several participants took part of the exercise over a period of 4 consecutive days. For the questionnaire: participants were all requested to fill up the questionnaires.

2.5.2 Background

The primary goal of October Demo was to determine if new tools and concepts developed under the JCDS 21 TDP helped the CF operators in achieving better operational effectiveness. Descriptions of these C2 applications developed under JCDS are included below:

a. Live Spaces: is an exploitation of the Australian LiveSpace technology to provide a smart meeting space supported by a smart room operating environment where hardware and software are integrated.

b. COPLanS: is an automated and distributed OPP tool including the latest decision support enhancements; time-sensitive planning, planning dependencies and implications management, analysis, risk management and reporting tools (including OPLAN, COPLAN, branch plan and sequel plans management tool).

c. K-Mapper: supports organizational management of knowledge assets and provides advanced link displays and reporting tools.

d. Advanced Command Portal: provides the foundations for a Command and Control Collaborative Environment (on top of Command View) supporting shared situation awareness, information management, systems integration and collaborative working.

e. Handheld Commander Tool: provides Commanders and senior staff officers on the move or with limited connectivity with real-time (or near real-time) access to, and interaction with, information from Command View and COPLanS in a secure manner.

f. Execution Manager: supports time-sensitive as well deliberative operations execution through continual automated monitoring of the situation inputs and execution reports.

g. Total Resource Visibility: is a decision support system for near real-time resource visibility providing asset information: identity, location, status, and condition of assets in the logistics chain.

The primary objectives of this event were to determine if the Integrated Command and Control Collaborative Environment (IC2CE) improves CF time-sensitive decision-making processes, improves shared SA within the Command Post (CP), and improves collaboration within the CP and with external agents.

Participants in October Demo assumed the role of JTFG in a scenario designed around the 2010 Olympic Games. Like Experiment 1, a primary characteristic of the scenario was its complexity, which can be defined as the composition and interaction of variables in a situation that affects the source and accuracy of information, the communication and decision-making processes, and the nature of the activities that are required in order to arrive at a desired end state. Complex military environments are characterized by a number of factors, for example time-sensitive response

activities, simultaneous life-threatening incidents across multiple geographic locations, multi-stakeholder interaction that require collaboration among organizational cultures to facilitate personnel and equipment resource sharing, information feeds and various communication channels, informal and formal processes, and the constraints of policy and legal issues. By their very definition, complex situations require the collaboration of a number of stakeholders. JTFG was, therefore, a suitable scenario to simulate complex situations for CF HQ personnel.

Like Experiment 1, this second experiment was run using the DRDC Valcartier Mobile Lab equipped with new hardware and software and meant to emulate an operation centre. October Demo took place over a 4 day period (21 – 23 October 2008), including 2 day of training and 2 mornings (approximately 4 hours/day) of experimentation.

Findings from October Demo was used to support Major Events Coordinated Security Solutions (MECSS). The MECSS project is taking steps to create collaboration across the federal Science and Technology (S&T) communities in order to support major events like the 2010 Vancouver Olympics.

2.5.3 Data Collection Plan

The DCP used in the October Demo was almost the same (only few minor modifications) than the one used during JCDS 21 EXP 1. The same metrics were used.

2.5.3.1 Observation team

The data collection team was made up of 8 people, including the Data Collection supervisors (Richard Breton and David Smith) who collected data as well.

Two data collectors were located in the Mobile Lab, 4 were located in the White Cell, and 2 LiveSpaces (see Table 7). Each was assigned specific responsibilities based on their role.

Table 7: Data Collection Team

Name	Organization	Experiment Location	Task Assigned
Richard Breton	DRDC Valcartier	White Cell	Supervisor
David Smith	DRDC Toronto	White Cell	Supervisor
TBD	Star Top	White Cell	Communications flows
TBD	TBD	White Cell	Communication flows
Marie-Eve Jobidon	DRDC Toronto	Mobile Lab	Communication flows

Elaine Maceda	DRDC Toronto	Mobile Lab	Communication flows
TBD	DRDC Toronto	Live Spaces	Communication flows
TBD	DRDC CORA	Live Spaces	Communication flows

2.5.3.2 Personnel assignment (Roles and Responsibilities)

All data collectors received comprehensive training on the C2 applications on Tuesday, which gave them the opportunity to get familiarized with the environmental context and the C2 applications. They also attended the training of the participants on the C2 applications. During participant training it is their responsibility to gather observations of the participants in order to familiarize themselves with the participants and their respective roles for the JTFG as well as with the group dynamic.

Two members of the JCDS 21 Data Collection team were assigned to the tasks associated with evaluating C2 applications and C2 processes. During Live Game Play (Wednesday and Thursday), these team members were located in the white cell and observed the task execution via a camera located in the Mobile Lab. This camera transmitted the visual and auditory feeds from the Mobile Lab to the observers' laptop. Headsets were used to accurately hear the verbal exchanges between the participants.

Two members of the JCDS 21 Data Collection team (Marie-Eve Jobidon and Elaine Maceda) were located in the Mobile Lab. Their core responsibilities included observing and documenting the communication and collaboration of the participants. They were equipped with a matrix that helped them categorize their observations of the participants' communication and collaboration activities during each experimental session.

Two members of the JCDS 21 Data Collection team were located in LiveSpaces. Their core responsibilities included observing and documenting the communication and collaboration of the participants. They were equipped with a matrix that helped categorizing their observations of the participants' communication and collaboration activities during each experimental session.

All data collection team members were required to re-transcribe their notes following each daily session in order to make their observations available as soon as possible. These notes were a vital source of data and were required to be as descriptive as possible, be presented in a bullet format (Word document), and be produced daily.

Richard Breton ensured that all questionnaire data was collected either through MORAE (or paper if need be). As such, he managed the questionnaires administration process (from MORAE) during the Live Game Play (Wednesday and Thursday).

2.5.4 Lessons learned from the October Demo

There are several similarities between the JCDS 21 EXP1 and the second JCDS 21 TD event, the October Demo. Both were events sponsored by the TD contrarily to the others preceding where the JCDS 21 TD team leveraged from these events and took the opportunity to refine its data collection, to showcase its applications and to develop its metrics. Consequently, the JCDS 21 EXP1 and October Demo used the same DCP and the same scenario.

However, the objective of the October Demo was different than the one pursued during the JCDS 21 EXP1. One objective was to determine if the IC2CE improves CF time-sensitive decision-making processes, improves shared SA within the CP, and improves collaboration within the CP and with external agents. However, another objective was to showcase the applications developed within the TD to potential users.

The importance of the demonstration sessions during the October Demo event influenced the quality of the data collection (observations and questionnaires) during the event.

2.5.4.1 Lessons Learned related to the metrics

As mentioned previously, the metrics used for the October Demo were exactly the same as those used during the JCDS 21 EXP1. Consequently, all lessons learned based on the administration of these metrics are presented in the preceding sections related to the JCDS 21 EXP 1.

2.5.4.2 Lessons Learned related to the level of realism

Observation: One major distinction between the October Demo and the JCDS 21 EXP1 events is the nature of the event. While the former was seen as a demonstration session, the latter was closer to an experiment. This distinction did not impact the nature of metrics used to collect data (same observation and questionnaire administration processes) but it did considerably influence the level of realism of the event. It is critical for an experiment to be as realistic as possible in order to collect data than can be applied in real life situations. In fact, the experimenters want the experiment to look as much as possible to real life situations. To do so, the interaction between the observers and experimenters must be as minimal as possible with the participants.

During demonstrations, the interaction between the observers and the participants is much more important. The observers are trained to interview the participant while doing their tasks in order to collect information about the systems used and their performance. As a result, the level of realism of the event is reduced.

October Demo was not an experiment. Consequently, it was not designed to *uniquely* investigate cause-and-effect relationships. The experimenters throughout the observation and questionnaires administration processes tried to determine the impact (their effect on the user performance) of the potential C2 applications. However, they wanted also to collect feedback for future design.

Recommendation: Taken as a demonstration session, the October Demo was a success for the JCDS 21 TD team. However, too many extraneous variables (i.e. presence of VIPs, several interventions of the observers with the participants) challenged the level of realism of the event

and then challenged the validity of the questionnaires results. Questionnaires may not be appropriate as a data collection strategy for such events.

2.5.4.3 Lessons Learned related to the level of availability of the participant

For a demonstration session, a key component is the availability of the participant during the session. During October Demo, several participants took part of the event. Their participation helped the JCDS 21 TD team to:

- Demonstrate the capability of the C2 applications to potential users; and
- To collect information and to gather feedback for future design cycles.

Observation: Taken as an experiment, one problem challenged the validity of the questionnaires results. During the event, different people played a given position throughout the event. Consequently, it was difficult to compare the results from day to day when these results reflected the answers of different people.

Recommendation: In the case of demonstration, questionnaires should be oriented to collect the participants' feedback instead of their performance or their level of awareness.

3 General guidelines

The objective of this section is to provide a set of guidelines to support the development of measurement protocols that include the selection of metrics and the establishment of a required level of realism and the availability of participants.

Depending of the nature of the event, some metrics are more appropriate for particular types of event than other. Characteristics of events are setting boundaries that should influence the selection of appropriate metrics. Among others, the most predominant characteristics are the level of realism required for the event, the level of intrusiveness tolerated in the environment for the application of the metrics, the number of participants and their accessibility. Based on this, we prepared the following guidelines.

3.1 Metric selection guidelines

Establishing the objective of the event

- **To train people**: the objective is to train actual or future operators for the execution of a specific task. Consequently, this activity may not be seen as scientific activity. Instead, the goal of this activity is to bring novices or beginners to an acceptable level of performance before being deployed in operational environments. While the primary nature of this activity may not be scientific, the application of measures is still required. In order to establish the distinction between beginners and experts, the performance typical of beginners and experts must be defined and measured.
  - Constraint: An important constraint for the application of a measure is the relatively high level of realism of the testing environment that is required.
- **To demonstrate technological applications**: the objective is to showcase technological applications to potential users in order to show their benefits related to a given task execution or to collect feedback (prototyping) to influence future system iterations. The objective of demonstration is to showcase future applications, and then it is not purely scientific. Prototyping is part of a scientific approach. A relationship between the feedback gathered and specific components of the systems must be established.
  - Constraint: The level of realism required for demonstration or prototyping sessions is still relatively high and may prevent the use of specific types of measures that require a certain level of intrusiveness in the experimental setting.
- **To understand task execution or the role of cognitive processes**: the objective is to establish a cause-to-effect relationship between a factor and an observed performance. Experiments are purely scientific. Hypotheses, based on scientific background, must be stated, measured, verified and validated.
  - Constraint: An important constraint is related to the number of participants required in order to reach an acceptable statistical power to verify the hypotheses.

- Constraint: Some experimental protocols require repeated measures (doing the experiment several times but varying the conditions in a controlled manner) in order to get several measures of the same concept in different conditions. This may be time and resources consuming and could be problematic when military personnel are required.

Interestingly, these three objectives are related to the three types of event executed in the context of the JCDS 21 TD project and documented in this report.

In the context of training it is important to identify activities or tasks that must be trained. In a demonstration, it is important to define properly the task and its subcomponents that should benefit from the new technological systems. In an experiment, the factors of interest should be clearly defined in order to understand a given effect and to relate it to a given cause (factor),. Then, a formal operational and/or scientific definition of that concept must be provided. It is critical that all people involved in the measurement activities work from the same definition of the concept.

For instance, situation awareness is a concept. Such concept should include different distinct factors (for instance the level of awareness, stress, fatigue, etc.). Thus, an exact definition of that concept is required in order to understand the impact of a given factor on that concept. The impact would be measured by differences in the human performance when the level of the factor is varied (for instance, manipulation the level of stress in the situation).

In order to measure properly a concept or the influence of a given factor on the performance, it is critical to clearly define this concept. In fact, the identification of the metrics used to measure the concept is dependent of the definition of the concept.

Defining the to-be measured concept

- **A specific task performance** (i.e. Situation Awareness or Decision Making): The concept to-be-measured can be a specific task performance. In that situation, the objective is to get a performance level about the task execution. However, it is critical that the measurement selection lays on a very precision task description in order to make sure that the performance measured represents specifically the performance of the task of interest.
  - Constraint: The study of specific task execution requires the reproduction of the conditions in which the task is usually performed in its operational environment. Consequently, it may ask for high-level simulators or field trials.
- **A specific cognitive process** supporting a task performance (i.e. perception, comprehension, selection, projection): In some measurement protocols, the objective is to study cognitive process in order to get knowledge in regards to their functioning. Here again, a complete and precise description of the cognitive process based on cognitive theories and models is essential. Such theories and models would provide background for hypotheses statement and to understand the observed results in regards to the studied cognitive process.
  - Constraint: The study of a cognitive process requires very controlled conditions in order to reduce as much as possible the impact of extraneous variables.

Consequently, depending on the complexity of the studied cognitive process, its study may only be possible in laboratory thus reducing its ecological validity.

- The effect of a **specific factor** on a task performance (i.e. time pressure, uncertainty, fatigue, stress): The objective of a testing and measurement activity can be the evaluation of the impact of a given factor and a given task performance. In order to relate the task performance (effect) and the impact of the factor (cause), a definition of the factor and its potential effect (stated as hypotheses) is required.
  - Constraint: The study of the impact of factor requires a systematic manipulation of this factor. Some ethical consideration may come into play in such experimental protocols.

Once the to-be-measured concept and the context (event objective) in which it will be measured are defined, the importance of the measurement activity, where it should take place and what is the window of opportunity to apply the metric during the military event, must be set. Military events, because of the time and resources involved, are often multi-purpose. Consequently, testing and measurement events must, most of the time, comply with important conditions defined by the nature of the military event. This condition defines the remaining guidelines.

Establishing the importance of the measurement activities within the military event

- **Secondary importance/several constraints:** In some circumstances, the military events provide opportunities to collect data but with several constraints.
  - Constraint: Measurement protocol must be compatible and compliant to the primary military objective.
  - Constraint: The application of metrics must be as unobtrusive as possible.
  - Constraint: Participants cannot be disturbed at any time during the military event.
- **Secondary importance/some constraints**: In other circumstance, even if the objective of the measurement activity is secondary in the overall military event, metrics can still be applied.
  - Constraint: Measurement protocol must be compatible and compliant to the primary military objective.
- **Primary objective/minimal constraints**: In this situation, the primary objective of the military event is to measure and to get a performance evaluation.
  - Constraint: the composition of the group of participants (homogeneity of the group, level of expertise) and their availability may become an importance issue.

Establishing the context in which the measurement activities take place

- **Field trials:** Field trials are extremely costly in terms of time and resources. Consequently, they are mostly used for important training events. They often occur at the end of a training program. Measurement can occur during these events in order to get an evaluation of the participants' performance right before their deployment in the operational environment.

- Constraint: Field trials require a very high level of realism that constraints significantly the type of measure that can be used.

- **In simulator:** Simulators are used for training, demonstration or experiment purpose. The objective with simulators is to recreate as much as possible the operational conditions while keeping the control over the variables of interest at an acceptable level.
  - Constraint: Simulators can be very costly and may require important technological setups.
- **In laboratory:** Laboratory settings are used to scientifically investigate specific cognitive processes. The level of realism in the experimental setup is compromised in favour of more control over the variables measured.
  - Constraint: It may be difficult to extrapolate results observed in those settings to operational environments.

Establishing the timing of the measurement activities (Window of opportunity)

- **Pre-event**: Measurement preceding an event (training, demonstration, experiment) can be done in order to get a baseline (comparative measure). In the case of training, it would establish the level of the participant before the training session. For demonstration, it could determine the actual performance of user (with their actual systems) in order to get to a comparative baseline. In experiment, pre-event measurement is often performed in order to get a baseline before the manipulation of variables of interest.
  - Constraint: It is critical to ensure that the any pre-event measurement does not provide any clues about the task to be executed during the event.
- **During the event:** Measurement during the event is essential to get an on-line evaluation of the participants' performance. Such measurement activities include observations and on-line data recording (time reaction, keystrokes, mouse-clicking, eye trackers, response selection, etc.).
  - Constraint: The level of intrusiveness allowed in the event is a key factor determining the type of metrics used.
- **Post-event:** Post-event measures consist in interviews and questionnaires administration. They are mostly used when on-line measurement is not feasible.
  - Constraint: Data recorded from interviews and questionnaires represent the content of participants' long-term memory and may be contaminated by other declarative knowledge (expertise and background) as observed by Rousseau et al [17].

The next Figure represents the five guidelines defined from the lessons learned during the JCDS 21 TD project. These guidelines should be used to define measurement protocols appropriate for the type of events (training, demonstration, experiment) occurring in a military environments.

;

Figure 2: Guidelines to select metrics for measurement in military events.

The five guidelines stated above define boundaries to support the selection of metrics based on the conditions prevalent in the testing and measurement event. Metrics can be divided into two general types (objective versus subjective). Subjective measures include questionnaires, interviews and observations. They are the beliefs or judgements of the participants (participants provide a self-evaluation in a questionnaire or during an interview) or a third person (for instance, the observations made by a SME). Objective measures are called event-based or performance measures. Their collection is not based on self-evaluation. They are related to different task parameters (time to react to a probe, selection, etc.). Their collection requires for technological support, for instance eye trackers, voice recording, keystrokes logs, etc.

3.1.1 Evaluation of potential metrics related to each guideline

The lessons learned during the five JCDS 21 TD events led to the identification of five general guidelines to support the selection of appropriate metrics. The following tables present a spreadsheet for each guideline that includes:

- Potential options in the context of military events. While the list may not be exhaustive, the included options represents to most generic ones.
- The most predominant constraints related to each option.
- Potential measures that could be used with each option
  - Constraints (in red) related to the use of each metrics.

Table 8: Establishing the objective of the event (First Principle).

Options	Important constraint	Potential measures and comments
To train people	Requires the reproduction of the task and the environment in which the task will be performed in the operational environment	-observation: Observation can be used to collect data about the quality of the training and the level of performance of participants. However, specific behaviours related to the quality of training and the level of performance must be observable (overt behaviours). Note: Data recorded are based on the evaluation of a tierce person (trained observers). To get valid data, several observers should record the same observation (inter-judge agreement). -questionnaire/interview: Questionnaires and interviews are a useful method to get information regarding the quality of the training program. However, it may be less efficient to get information regarding the participants' level of performance. Note: Such data collection strategy would require that the participants are self-evaluating their own performance. Then, results may reflect more the participant's feeling about their performance than the performance itself. Also, to get valid data, instructions should be clearer as possible to make sure that the evaluation is made on what occurred during the training session instead of based on declarative knowledge (expertise and background). -objective measure: Objective measure is a very interesting way to measure the impact of a training program. It could provide different level of performance at different moment in time. For instance, a task performance level could be based on a time to react of a threat. A significant reduction of the time to react throughout the training session could be an indicator of a performance improvement. Note: Because of the importance to recreate the conditions in which the task is executed for valid training and the importance of the technological support required to collect the data, the level of obtrusiveness of metrics must be considered.
To demonstrate technological applications	Requires the reproduction of the task that should benefit from the new technology	-observation: Observation can be used to collect data about a specific prototype (in the case of a prototyping session) or to evaluate the impact of a potential new device. However, specific behaviours defining any performance improvement must be determined and be observable (overt behaviours). Note: Data recorded are based on the evaluation of a tierce person (trained observers). To get valid data, several observers should record the same observation (inter-judge agreement). -questionnaire/interview: Questionnaires and interviews are a useful method to get information about the participants' feelings over a new application (level of trust, acceptance, comfort, etc.) or to gather feedback to influence future prototypes. Note: The evaluation of future technological applications through questionnaires and interviews should reflect the participant's feelings about the potential devices than the actual participants'

		level of performance with the new devices. -objective measure: Objective measure is a very interesting way to measure the potential impact of a new device on a task performance. For instance, the measurement protocol should be design to provide performance indicators when participants are using the new systems and when they use the actual ones. Note: Prototyping and demonstration require a certain level of environment realism. Then, the level of obtrusiveness of metrics must be considered.
To understand task execution or the role of cognitive processes	Requires several participants and repeated measures in order to reach statistical validity	-observation: Observations can be used to understand a given task execution if this task execution is mainly composed of overt behaviours. Note: However, it is not applicable to study the role of a given cognitive processes that are not observable (for instance, thinking, deciding, perceiving, etc.). -questionnaire/interview: Questionnaires and interviews can be used to understand a given task execution if this task execution is mainly composed of overt behaviours and the participants are aware of these observable steps. Note: Such measurement strategy is less applicable for evaluating the role of a given cognitive process in a task execution. It would require that participants are fully aware of the cognitive processes involved in their task execution. -objective measure: Objective measures may be the best way to evaluate a task execution or the role of a cognitive process on a given task execution. Note: The study of a task execution may require the reproduction of the environmental conditions in which the task is usually performed. Thus, the level of obtrusiveness of metrics that can challenge the realism of the situation must be considered.

Table 9: Defining the to-be measured concept (Second Principle).

Options	Important constraint	Potential measures and comments
The Execution of a cognitive task (i.e. Decision-Making or Situation Awareness)	Requires the reproduction of the task under study at high level of realism	-observation: If the task is composed of observable behaviours, the evaluation of the task execution is possible from observations. Note: This would require trained observers or SME in the environment that could challenge the level of realism of the situation. Also, data recorded are based on the evaluation of a tierce person (observers). To get valid data, several observers should record the same observation (inter-judge agreement). -questionnaire/interview: The evaluation of the task execution is possible if participants are aware of the cognitive steps (not necessarily observable) or overt behaviours required executing a task. For instance, it may be difficult for participants to list precisely the cognitive steps such as perception, comprehension, etc. It could be easier to list observable steps such as select a

		track, engage a track, etc. Note: Data recorded are based on self-evaluation from participants. To get valid data, instructions should be clearer as possible to make sure that the evaluation is made on what occurred during the event instead of based on declarative knowledge (expertise and background). -objective measure: Objective measures may be the best way to evaluate a task execution. It does not rely on subjective self-assessment or the evaluation of a tierce person (the observers). Objective measures are based on the measurement of parameters related to the task performance such as time reaction to a probe, selection, keystrokes, eye movements, etc. Note: Because of the importance to recreate the conditions in which the task is executed and the importance of the technological support required to collect the data, the level of obtrusiveness of metrics must be considered. Also, objective measures require rigorous experimental protocols in which controlled or measured variables are manipulated on different conditions. Consequently, it requires several measures of the same variable.
The role of a cognitive process in a task execution (i.e. comprehension, perception, attention)	Requires controlled conditions	-observation: Not applicable. Cognitive processes (thinking, deciding, perceiving, etc.) are not observable. -questionnaire/interview: Hardly applicable. Using questionnaires of interview sessions would require that participants are fully aware of the cognitive processes involved in their task execution. -objective measure: Objective measures are surely the best way to evaluate the role of a cognitive process in a task execution. Objective measures are based on the measurement of parameters related to the task performance such as time reaction to a probe, selection, keystrokes, eye movements, etc. These parameters are defined from cognitive theories and models that also support the statement of hypotheses. Note: Objective measures require rigorous experimental protocols in which controlled or measured variables are manipulated on different conditions. Consequently, it requires several measures of the same variable.
The study of the impact of a Human/Environmental Factor on a task execution (i.e. Human: stress, fatigue; Environmental: time constraint)	Requires systematic factor manipulation	-observation: the study of the impact of environmental factor (effect of time constraint, uncertainty, etc.) may be appropriate for a data collection strategy based on observation than the impact of human factor (effect of stress, fatigue, etc.). Note: Data recorded are based on the evaluation of a tierce person (observers). To get valid data, several observers should record the same observation (inter-judge agreement). -questionnaire/interview: The impact of human/environmental factors on a task execution can be evaluated from questionnaire or throughout interview sessions. Note: Data recorded are based on self-evaluation from participants. To get valid data, instructions should be clearer as possible to make sure that the evaluation is made on what occurred during the event instead of based on declarative knowledge (expertise and background). -objective measure: Objective measures provide a non-biased measure of a phenomenon. Then, when possible, such

		measurement strategy should be favoured to measure the impact of an external or internal factor on a task performance. Note: Objective measures require rigorous experimental protocols in which controlled or measured variables are manipulated (for instance, various level of time constraints or fatigue) on different conditions. Consequently, it requires several measures of the same variable.

Table 10: Establishing the importance of the measurement activities within the military event (Third Principle).

Options	Important constraint	Potential measures and comments
Secondary/Several constraints	Measurement protocol must be compatible and compliant to the primary military objective. The application of metrics must be as unobtrusive as possible and the participants cannot be disturbed at any time during the military event.	-observation: Observations may be the best way to collect data when important constraints such as the availability, level of realism and the need to get unobtrusive measures come into play. Note: It is important to negotiate with stakeholders the presence of enough observers in the participants' environment. Not enough observers by participants (very low ratio) would lead to very few observations. -questionnaire/interview: Note: Questionnaires and interviews should be difficult to administrate and run when the measurement activities are secondary and the conditions are very constraining. -objective measure: Objective measures may be appropriate to get metrics as unobtrusive as possible. Note: Objective measures require important technological settings that should be compatible with the environment in which the measurement occurs.
Secondary/some constraints	Measurement protocol must be compatible and compliant to the primary military objective.	-observation: Observations may be the best way to collect data when the data collection must be included within larger military event with non-related primary objectives. Note: It is important to negotiate with stakeholders the presence of enough observers in the participants' environment. Not enough observers by participants (very low ratio) would lead to very few observations. -questionnaire/interview: The administration of questionnaires and interviews could be possible if there are windows of opportunity for such data collection activities in the military event. Note: It is important that such data collection activities being compatible with the primary objective of the event. -objective measure: Objective measures may be appropriate to get metrics as unobtrusive as possible. In that sense, they should not disturb the military event. Note: Objective measures require important technological settings that should be compatible with the environment in which the measurement occurs.
Primary/minimal	The composition of the group	-observation: Observations are possible specifically if the

constraints	of participants (homogeneity of the group, level of expertise) and their availability may become an importance issue.	participant availability/accessibility is restricted. Note: It is important to optimally position the observers in the environment in order to make sure that key positions are covered (if the homogeneity of the group is low: some positions more important than others). -questionnaire/interview: The administration of questionnaires and interviews could be possible if there are windows of opportunity for such data collection activities in the military event. Note: It is important that all positions are represented if the homogeneity of the group is an issue. -objective measure: Objective measures may be the best way to collect data if the primary objective of the activity is to collect data. In that sense, resources and efforts can be deployed to develop experimental setups that allow such more complex but fruitful data collection strategy.

Table 11: Establishing the context in which the measurement activities take place (Fourth Principle).

Options	Important constraint	Potential measures and comments
During field trials	Requires very important technological and environmental considerations	-observation: Observations during field trials are possible if the observers are allowed in the operators' environment. Remote observations can be an alternative to support such data collection strategy when the physical operator's environment does not provide the opportunity to host an observer (for instance, in a F-18 cockpit). Note: An optimal ratio between the number of participants and the number of observers should be defined. Such ratio should consider the nature of the task (for instance, tempo: high versus low), the distinctiveness of the behaviours to observe, the level of intrusiveness acceptable in the situation, etc. -questionnaire/interview: Note: The administration of questionnaires and interviews during field trials may not be possible specifically during training or demonstration sessions in which a certain level of realism is required. -objective measure: Objective measures could be used in field trials if the environmental conditions of the field trials allow for such strategy. Note: Field trials may require the reproduction of the environmental conditions in which the task is usually performed. Thus, the level of obtrusiveness of metrics that can challenge the realism of the situation must be considered.
In simulator	Requires important technological setup to simulate the operational environment	-observation: Observations within a simulator are possible if the observers are allowed in the simulator. Remote observations can be an alternative to support such data collection strategy when the physical operator's environment does not provide the opportunity to host an observer (for instance, in a F-18 simulator cockpit). Note: An optimal ratio between the number of

		participants and the number of observers should be defined. Such ratio should consider the nature of the task (for instance, tempo: high versus low), the distinctiveness of the behaviours to observe, the level of intrusiveness acceptable in the situation, etc. -questionnaire/interview: Note: Simulators are used to simulate the environmental conditions in which a task is executed at a fairly high level of realism. Consequently, the administration of questionnaires or interviews sessions should be avoided. -objective measure: Simulator should provide great opportunities to collect objective measures. Note: The technological setup required to collect the data should be compatible with the simulator technological setup.
In laboratory	Requires technological setup to control the experiment (control the variable, the scenario, etc.).	-observation: Observations are possible in a laboratory. Note: A laboratory setup is often chosen for its capacity to restraint the impact of extraneous variable in the experiment. The presence of observers can be seen as an extraneous variable. -questionnaire/interview: Note: Laboratory setups are used to create very controlled conditions in order to make sure that the observed results are related to the cause without any external interference. Consequently, the administration of questionnaires or interviews sessions should be avoided. -objective measure: Laboratory setups include most of the time objective measures are data collection strategy. Such strategy is compatible with the needs of control and the reduction of the impact of extraneous variables. Note: The technological setup required to collect the data should be compatible with the laboratory technological setup.

Table 12: Establishing the timing of the measurement activities (Fifth Principle).

Options	Important constraint	Potential measures and comments
Pre-event	Questionnaire should not provide information on the task performed during the event	-observation: Not applicable. -questionnaire/interview: Questionnaires and interviews can be used to get a performance baseline from which any other performances (i.e. following a training program, when using new devices, etc.) can be compared. -objective measure: Not applicable.
During the event (online)	Metrics should be selected based on the acceptable level of metric intrusiveness	-observation: Observations can be collected during the live play. Note: An optimal ratio between the number of participants and the number of observers should be defined. Such ratio should consider the nature of the task (for instance, tempo: high versus low), the distinctiveness of the behaviours to observe, the level of

		intrusiveness acceptable in the situation, etc. -questionnaire/interview: Note: The administration of questionnaires and interviews during live play may not be possible specifically in high tempo situations and situations requiring a certain level of realism. Such activities could disrupt the live play. -objective measure: Objective measures may be the best way to collect data online. They are not obtrusive and are collected without the participants being aware of the data collection process. Note: The technological setup required to collect the data should be compatible with the operators' technological environment.
Post-event	Results could reflect the content of participants long-term memory (knowledge) instead of what they experienced during the event	-observation: Not applicable. -questionnaire/interview: Questionnaires and interviews are the only data collection method to get information at the end of a live play. -objective measure: Not applicable.

3.2 Advantages and constraints for each type of metrics in the context of military events

This section summarizes the requirements, advantages, inconvenient for each type of metric (observation; questionnaire; interview and objective measures) in the context of military events (training, demonstration, experiment).

3.2.1 Observation

Requirements:

- trained observers or SME
- observable behaviours (physical, verbal)
- set of targeted behaviours
- an optimal ratio between observers and participants

Advantages:

- The intrusiveness of the metrics is only related to the presence of an observers in the room.
- Provide a good alternative when the participants' accessibility is low.
- Provide a direct evaluation of a given behaviour within an overall task performance.

Inconvenient:

- Based on the evaluation of a tierce person.
- May require a too high ratio of observers/participants.
- Observers should be closer as possible of the action to record all interesting data.

Best suited for (in order of preference):

1. Training events in Field trials
2. Training events in simulators
3. Demonstration in Field trials or simulators

3.2.2 Questionnaire

Requirements:

- Standardized and validated questionnaire
- Participants accessibility
- Task execution that allows self-evaluation
- Administration periods (pre-event; during and post-event)

Advantages:

- Not intrusive when used before and after the event.
- Can be used with several participants.
- Rapid data collection and analyzed (if the questionnaire is administrated electronically).

Inconvenient:

- Based on the participant self-evaluation (biased).
- The participants understanding of the questions and the way they answer to the questions may fluctuate from one participant to another.

Best suited for (in order of preference):

1. Prototyping
2. Demonstration in Field trials or simulators (administration before and after the event)
3. Training events in Field trials (administration before and after the event)
4. Training events in simulators (administration before and after the event)

3.2.3 Interview

Requirements:

- Requires one-to-one meeting with participants
- Participants accessibility
- Task execution that allows self-evaluation

Advantages:

- Not intrusive when executed before and after the event.
- Can provided very interesting feedback.

Inconvenient:

- Based on the participant self-evaluation (biased).
- The interview is time and resources consuming.
- The data analysis is time and resources consuming.
- The participants understanding of the questions and the way they answer to the questions may fluctuate from one participant to another.
- Dependent of the quality of the interviewer and the interaction between the interviewer and the interviewee.

Best suited for (in order of preference):

1. Prototyping
2. Demonstration in Field trials or simulators (meeting before and after the event)
3. Training events in Field trials (meeting before and after the event)
4. Training events in simulators (meeting before and after the event)

3.2.4 Objective measures

Requirements:

- Data collection requires technological support
- Data analysis requires technological support

Advantages:

- Can be done without interacting directly with the participants. The participants may not be aware of the data collection process.
- Data unbiased by participants' self-evaluation of their performance.
- Can provided very large pool of data related to the participant's performance.

- The data analysis is relatively simple and rapid.

Inconvenient:

- It requires a minimal number of participants in order to provide valid data.
- It requires a technological setup that must be introduced into the participants' environment.

Best suited for (in order of preference):

1. Study of the role of cognitive process in a task execution in laboratory or simulator setups
2. Prototyping in simulator environments
3. Demonstration in simulator environments
4. Training events in simulators

4 Conclusion

The objectives of the JCDS 21 TD project were related to the understanding of complex C2 tasks executed in complex integrated C2 environments and the design and development of technological solutions to support these tasks in these environments. To support those objectives, performance measurement and feedback gathering were important parts of the scientific approach in the project.

Measurement in military environments is a very complex endeavour. Consequently, the five JCDS 21 TD events documented in this report were a great opportunity for the JCDS 21 TD team to understand the subtlety and complexity of measuring performance in military settings. The main objective of the JCDS 21 TD project was to support the design process of support systems, but, as a side benefit, the team had the opportunity to understand, from lesson learned, the complexity of the measurement process.

This report suggests five guidelines that should be used to select appropriate metrics and experimental protocols based on the constraints of the environment in which the metric will be used and the requirements of the potential metrics. In fact, the selection of metrics should be seen as a trade-off between environmental constraints and metrics requirements as depicted in Figure 3. The military environment offers opportunities and constraints that should eliminate certain types of metrics. On the other hand, each specific metric has a set of requirements for its optimal application.

Figure 3: Trade-off between environmental constraints and metrics requirements.

The benefit of this report is the identification of the trade-off between these constraints and requirements. The spreadsheets presented in the previous section are a good example of such trade-off analysis.

The five guidelines originated from the lessons learned should provide support for the selection of appropriate metrics for future military events. They raise aspects that need to be considered when planning and designing measurement activities. They address the objective behind the measurement activity; they stress the importance of a clear definition concept; they define the relative importance of the measurement activity within the larger military event; and they include environmental constraints such as the time and the environment setting in which the measurement activity should take place.

Some conclusions can be held from the five events (Ardent Sentry 06, Friendly Lance, LiveSpace demonstration, EX PG and October Demo) executed in the JCDS 21 TD project.

- It was easier to base the data collection strategy on observations;
  - the ratios between observers and participants were relatively low
- It was difficult to administrate questionnaires at many times during the event;
  - then, it was difficult to reach a sufficient statistical power when analysis the data
  - questionnaires were sometime not filled properly (instructions not enough clear)
- Interviews were only good to get participants feedback about potential technological systems;
- Objective measures were not favoured.

In the future, effort should be devoted to develop metrics and protocols based on objective measures. Objective measures present several benefits:

- technological setup not intrusive;
- data recording not obtrusive;
- direct performance measure (not biased by self-evaluation or evaluation from a tierce person);
- rapid to analysis;
- provide several measure.

Also, there is a need to develop an infrastructure that should support the data recording. Finally, a toolbox including metrics specifically developed for being used in military settings is required.

References

[1] "Guide for Understanding and Implementing Defence Experimentation (GUIDEx)", The Technical Cooperation Program (TTCP), February, 2006.

[2] Rasmussen, J., Pejtersen, A.M., & Goodstein, L.P. (1994). "Cognitive Systems Engineering", New-York: Wiley.

[3] Jobidon, M.E., Rousseau, R., & Breton, R. (2005). "The effect of variability in temporal information on the control of a dynamic task". Theoretical Issues in Ergonomic Science, 6(1), 13 pages.

[4] Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., & Jenkins, D.P. (2005). "Human Factors Methods: A Practical Guide for Engineering and Design", Ashgate Publishing Limited, England.

[5] Drury, C. (1990). "Methods for Direct Observation of Performance", in J.R. Wilson & E.N. Corlett (eds), Evaluation of Human Work, 2nd edition, London: Taylor and Francis, pp.45-68.

[6] Lafond, D., Tremblay, S., Rousseau, R., & Breton, R. (2008). "Metrics for Assessing Support of Cognition in Command and Control (MASC3): Project Overview and Demonstration of the S-CCS microworld". Poster presented at the 2nd International Conference on Applied Human Factors and Ergonomics, Las Vegas, NV, USA.

[7] Kirk, R.E. (1992). "Experimental Design: Procedures for the Behavioral Science, 2nd Edition". Brooks/Cole Publishing Company, California, USA.

[8] Proctor, R.W., & Van Zandt, T. (1994). "Human Factors in Simple and Complex Systems". Allyn and Bacon, Massachusetts, USA.

[9] Breton, R., Paradis, S., & Roy. J. (2002). "Command Decision Support Interface (CODSI) for Human Factors and Display Concept Validation". Proceeding of Fusion 2002, Annapolis, USA.

[10] Chapanis, A. (1996). "Human Factors in Systems Engineering", Wiley-Interscience Publication. USA.

[11] Alberts, D.S, & Hayes, R.E. (2005). "Campaigns of Experimentation: Pathways to Innovation and Transformation". DoD Command and Control Research Program, USA.

[12] Bélanger, M., Guitouni, A., & Pageau, N. (2009). "Decision support tools for the operational planning process". Proceeding of the 14th ICCRTS: C2 and Agility. Washington, DC, USA.

[13] Salas, E., Sims, D.E., & Burke, C.S. (2005). "Is there a 'big five' in teamwork"? Small Group Research, 36(5), 555-599.

[14] Lafond, D., Tremblay, S., Dubé, G., Rousseau, R., & Breton, R. (2010). “A method and tool for estimating the costs/benefits of teamwork in different C2 structures”. Proceeding of the 15th ICCRTS: The Evolution of C2, Santa Monica, CA, USA.

[15] MacMillan, J., Paley, M. J., Levchuk, Y. N., & Serfaty, D. (1999). “Designing the Information Space and Physical Layout for a Command Center Based on an Optimized Organizational Structure”. Paper presented at the Command and Control Research and Technology Symposium, Newport, RI.

[16] MacMillan, J., Serfaty, D., Paley, M. J., & Levchuk, Y. N. (2000, July). “Designing the Information Space for a Command Center Based on Team Performance Models.” Paper presented at the International Symposium on Optical Science and Tech, San Diego, CA.

[17] Rousseau, R., Tremblay, S., Banbury, S., Breton, R., & Guitouni, A. (printing). “The Role of Metacognition in the Relationship between Objective and Subjective Measures of Situation Awareness”. Theoretical Issues in Ergonomics Science.

Annex A ARDENT SENTRY (AS) 06 - PRE-EXERCISE SURVEY

Instructions

The purpose of this survey is to obtain background information with respect to your experience and training with the tools and web capabilities forecasted to be used during AS 06. All individuals participating in AS 06 are to complete both the Pre-Exercise and the Quick Look Survey. However, for those individuals who are not participating in AS 06, it would be appreciated if you can take the time to fill out the Pre-Exercise survey. All data gathered in this survey will be held in confidence (demographic data will be removed before it is posted), and will be used for analysis purposes only.

The following survey is to be completed by COB on ***xx May 06.***
If you have any difficulties accessing or submitting the survey on C-Net, please contact ????.

Demographic Data

Last Name: _____
Rank: _____
MOC: _____
Years of Service: _____
Organization: _____
Position/Role (at your current job): _____

Systems (C-Net Workstations)

Note: C-Net workstations refers to TITAN, MCOIN, AFCCIS and LFC2IS

1) Do you always have access to a C-Net workstation for your work?
   Yes ___
   No ___
   Do not know ___
   (a) If you answered 'No', please explain why (for example, connectivity issues, sharing a workstation with many other users, location of the workstation). _____

2) How satisfied are you with the technical support for each of following items? Select only one response for each.

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied	Never had problems
Command View	°	°	°	°	°	°
Chat	°	°	°	°	°	°
IMS	°	°	°	°	°	°
C-Net hardware and connectivity	°	°	°	°	°	°

Comments ____

Tools

Note: The tools being considered are Command View (CV), Unclassified CV, Chat, the Incident Management System (IMS), C2PC and the Alta Vista Search Engine.

3) Do you have an account for the following tools?

	Yes	No
SameTime Chat	°	°
IPWaR	°	°
IMS	°	°
C2PC	°	°

Comments ___

4) How often do you use the following tools? Select only one response for each tool.

	Every Day	4 to 6 times a week	2 to 3 times a week	Once a week	Less than once a week	Never
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
SameTime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS for viewing	°	°	°	°	°	°
IMS for posting	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
Altavista Search Engine	°	°	°	°	°	°

Comments ___

5) To what extent do you agree that the following tools are easy to use? Select only one response for each tool.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Not applicable
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
SameTime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS for viewing	°	°	°	°	°	°
IMS for posting	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
Altavista Search Engine	°	°	°	°	°	°

Comments ___

6) To what extent do you agree that the following tools are useful and relevant for your work? Select only one response for each tool.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Not Applicable
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
SameTime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
Altavista Search Engine	°	°	°	°	°	°

Comments ___

7) Indicate if you have had any of these types of training for the following tools;
   - Formal instruction (i.e. courses, tutorials),
   - Informal instruction (i.e. demonstration from co-worker),
   - On-the-job training (i.e. learned on your own).

Select all that apply.

	Formal	Informal	On-the-job	None	Not sure
Command View	°	°	°	°	°
Unclassified CV	°	°	°	°	°
SameTime Chat	°	°	°	°	°
IPWaR	°	°	°	°	°

IMS for viewing	°	°	°	°	°
IMS for posting	°	°	°	°	°
C2PC	°	°	°	°	°

Comments ___

8) How satisfied are you with the training that you have received for the following tools? Select only one response for each tool.

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied	Did not receive training
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
SameTime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS for viewing	°	°	°	°	°	°
IMS for posting	°	°	°	°	°	°
C2PC	°	°	°	°	°	°

Comments ___

Collaborative Planning

For Questions 12 and 13, to what extent do you agree with the following statements?

9) I feel comfortable participating in web-based collaborative planning:
   Strongly agree ___
   Agree ___
   Neutral ___
   Disagree ___
   Strongly disagree ___
   Comments ___

10) I support new initiatives in web-based collaborative planning.
   Strongly agree ___
   Agree ___
   Neutral ___
   Disagree ___
   Strongly disagree ___
   Comments ___

11) For each stage of the Operational Planning Process, to what extent do you agree that it can be conducted using the tools (i.e. Command View, IPWaR, Sametime Chat, IMS) currently available on C-Net?

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Do not know
Initiation	°	°	°	°	°	°
Orientation	°	°	°	°	°	°
Course of Action development	°	°	°	°	°	°
Plan Development	°	°	°	°	°	°
Plan Review	°	°	°	°	°	°

SOP

12) Which of the following documents have you read? Select all that apply.

Chat – Overall Guidance ___
Chat – Sametime (Annex A) ___
Chat – Distributed Collaborative Planning (DCP) Secure CONOPs ___
Chat – IPWaR Contacts ___
Joint Command System (JCS) SOP ___
Sametime Quick Start Guide ___
Did not read any of these documents yet ___

Comments ___

General

13) Are there any other comments that you would like to make that were not covered in the survey? ____

This page intentionally left blank.

Annex B ARDENT SENTRY (AS) 06 - QUICK LOOK SURVEY

Instructions

Please complete the Pre-Exercise Survey for AS 06 before starting this one. Everyone who participated in AS 06 is requested to complete both surveys.

The purpose of this survey is to capture your immediate impressions of AS 06. All information collected will be aggregated so that individual responses will not be distinguishable, and all individual responses will be kept confidential.

The following survey is to be completed when you have finished your participation in AS 06, and at the latest by COB ?, ***May 06***.

If you have any difficulties accessing or submitting the survey on C-Net, please contact ??????.

Demographic Data

Last Name: _____
Rank: _____
MOC: _____
Years of Service: _____
Organization: _____
Position/Role (during AS 06): _____
What days did you participate in AS 06?

Mon 08 May 06	Yes ____	No ___
Tue 09 May 06	Yes ____	No ___
Wed 10 May 06	Yes ____	No ___
Thu 11 May 06	Yes ____	No ___
Fri 12 May 06	Yes ____	No ___

Exchange of Operational and Strategic Level Information

Note: By definition, "tools" includes Command View (CV), Unclassified CV, the Incident Management System (IMS), Chat, C2PC, and DSEL.

1) For each of the following tools, to what extent do you agree that the tool was useful to help you gain situational awareness during AS 06? Select only one response for each tool.

Tools	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did not use tool
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
Sametime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
DSEL	°	°	°	°	°	°

Comments ____

2) How often did you exchange operational information (OI) (i.e., obtain and/or give OI) with each of the following organizations during AS 06? Select only one response for each organization.

Organizations	Never	1-3 times	4-6 times	7 or more
SJS	°	°	°	°
Canada COM	°	°	°	°
CEFCOM	°	°	°	°
CANSOFCOM	°	°	°	°
CANOSCOM	°	°	°	°
JTF-ATLANTIC	°	°	°	°
JTF-CENTRAL	°	°	°	°
N-NC	°	°	°	°
AOC	°	°	°	°
CDI	°	°	°	°
Other (please comment below)	°	°	°	°

Comments _____

3) How often did you use each of the following sources to find specific OI during AS 06? Select only one response for each source.

Sources	Never	1-3 times	4-6 times	7 or more
Command View	°	°	°	°
Unclassified CV	°	°	°	°
Sametime Chat	°	°	°	°
IPWaR	°	°	°	°
IMS	°	°	°	°
C2PC	°	°	°	°
GP Net	°	°	°	°
RELCAN	°	°	°	°
NIPRNET	°	°	°	°
Telephone	°	°	°	°

DSEL	°	°	°	°
STU III	°	°	°	°
Unclassified email	°	°	°	°
Classified email	°	°	°	°
Unclassified fax	°	°	°	°
Other (please comment below)	°	°	°	°

Comments _____

4) For each of the following web pages on CV, to what extent do you agree that the OI resident on that site was easy to find? Select only one response for each web page.

Web page	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Not used	Not applicable
Strategic	°	°	°	°	°	°	°
CanadaCOM	°	°	°	°	°	°	°
CEFCOM	°	°	°	°	°	°	°
CANSOFCOM	°	°	°	°	°	°	°
CANOSCOM	°	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°	°
JTF-ATLANTIC	°	°	°	°	°	°	°
JTF-CENTRAL	°	°	°	°	°	°	°
CDI	°	°	°	°	°	°	°
N-NC	°	°	°	°	°	°	°
AOC	°	°	°	°	°	°	°

Comments _____

5) If you visited the CDI web page, to what extent do you agree that the CDI web site was able to provide all your intelligence requirements? Select only one response.

Strongly Disagree	Disagree	Neutral	**Agree**	Strongly Agree	Do not know, N/A
°	°	°	°	°	°

Comments ____

6) To what extent do you agree that the OI you obtained during AS 06 was timely and relevant? Select one response for each criterion.

Information was:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Not applicable
Timely	°	°	°	°	°	°
Relevant	°	°	°	°	°	°

(a) If you tended to disagree, please comment on what information was not timely and/or not relevant, and where this information was obtained. _____

7) What additional information, if any, should be posted or linked to Command View?_____

8) List any other suggestions on how to improve Command View. ___

Posting and Editing Information

9) Do you have a login to post and/or edit information on the following systems?

System	Yes	No
C-Net	°	°
Unclassified	°	°

Comments _____

If you answered 'No' for both systems, please proceed to Question 12.

10) Were you required to post and/or edit information during AS 06?

System	Yes	No
C-Net	°	°
Unclassified	°	°

Comments: _____

11) Do you require additional training to post and/or edit information?
Yes ____ No ___
If you answered 'yes', please specify what kind of training you would find useful.
_____
__

Significant Incident Reporting

12) Did you report any significant incidents?

Yes ___ No ___ Do not know___

(a) If you answered 'Yes', did you use specific SOP for incident reporting?
Yes ___
No ___

Do not know___

(b) If you answered “Yes”, please comment _____

13) To your knowledge, were there any problems/difficulties related to reporting significant incidents during AS 06?

Yes ____

No ____

If you answered “Yes”, please comment _____

Recognized Air Picture (RAP)

14) During AS 06, were you required to find information about military air resources?

Yes ___ No ___

(a) If you answered ‘Yes’, to what extent do you agree that the following tools were useful to find the required information? Select only one response for each tool.

Tools	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did not use tool	Not applicable
Command View	°	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°	°
Sametime Chat	°	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°	°
IMS	°	°	°	°	°	°	°
C2PC	°	°	°	°	°	°	°
DSEL	°	°	°	°	°	°	°
Other (please comment below)	°	°	°	°	°	°	°

Comments ___

15) List any suggestions on how to improve the RAP. __

Training, SOP and System development

16) For each of the following items, to what extent do you agree that more training prior to AS 06 would have helped your performance during the exercise? Select only one response for each item.

Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Do not know
IMS	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
Sametime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
Command View	°	°	°	°	°	°
Command View	°	°	°	°	°	°
DSEL	°	°	°	°	°	°
Significant Incident Reporting	°	°	°	°	°	°
Operational Planning Process (OPP)	°	°	°	°	°	°

Comments ____

17) List any other topics (if any) that would be useful to include in a training program for future Command and Control exercises. _____

18) For each of the following SOP, to what extent do you agree that they are useful as currently written? Select only one response for each SOP.

SOP	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did not read
Chat (Overall Guidance)	°	°	°	°	°	°
Chat – Sametime (Annex A)	°	°	°	°	°	°
Chat – Distributed Collaborative Planning (DCP) Secure CONOPS	°	°	°	°	°	°
Chat – IPWaR Contacts	°	°	°	°	°	°
Joint Command System (JCS) SOP	°	°	°	°	°	°
SameTime: Quick Start Guide	°	°	°	°	°	°

(a) If you tend to disagree, please explain why. ____

19) List any topics (if any) for which SOP need to be developed. _____

20) Did you experience any technical difficulty (e.g. bugs, slow response, inactive links, etc.) with the following tools? Select only one response for each tool.

Tools	Yes	No	Did not use tool
Command View	°	°	°
Unclassified CV	°	°	°
Sametime Chat	°	°	°
IPWaR	°	°	°
IMS	°	°	°

C2PC	°	°	°
DSEL	°	°	°

(a) If "Yes, please specify. ___

21) To what extent do you agree that the following tools facilitate your work? Select only one response for each tool.

Tools	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did not use tool
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
Sametime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
DSEL	°	°	°	°	°	°

Comments ___

22) For each of the following tools, to what extent do you agree that more improvements are required in order to help you do your job? Select only one response for each tool.

Tools	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Did not use tool
Command View	°	°	°	°	°	°
Unclassified CV	°	°	°	°	°	°
Sametime Chat	°	°	°	°	°	°
IPWaR	°	°	°	°	°	°
IMS	°	°	°	°	°	°
C2PC	°	°	°	°	°	°
DSEL	°	°	°	°	°	°

Comments ___

23) What improvements to the tools would you like to see? _____

CF Joint Operational Planning Process (OPP)

24) Did you participate in operational planning during AS 06?
Yes (please specify) ____
No ____
Do not know ____

Comments _____

General

25) What were the strengths of AS 06? ___

26) What areas in the design or execution of the ARDENT SENTRY exercise need to be improved for next year? ___

27) Any final comments? ____

THANK YOU!

Annex C NATO TG-127 project on Military Command Team Effectiveness

Effective teamwork is recognized as a critical mission success factor. NATO Task Group - 127 has been working to support commanders in assessing and improving the performance of their command or staff teams. We have developed a Command Team Effectiveness support tool that permits commanders and team members to assess their teams' effectiveness. The tool is based on an analysis of the literature and practical command experiences. We are now attempting to validate the tool by applying it in a variety of training and operational settings. We kindly request your contribution to this project, which would be of great value to all NATO countries and their commanders. If you have any questions about this research project, then please contact a member of the project team.

PROJECT TEAM:
EUROPE: peter.essens@tno.nl Dr. P.J.M.D. Essens, TNO Human Factors, NL.
USA: jay.goodwin@hqda.army.mil Dr. J. Goodwin, US Army Research Institute, USA.
CANADA: joe.baranski@drdc-rddc.gc.ca Dr. J.V. Baranski, Defence R&D Canada, CA.

Instructions

The instrument consists of several parts. The main part asks you to assess a number of items that are related to your team: the mission, the task, the organisation, the team members, the team behaviours, and the outcomes. We would like you to assess each item four times:

a. Was the item relevant for your team (relevance)?
b. To what extent was the item present in your team (quality/magnitude)?
c. Did the given quality/magnitude have a positive, negative, or no influence on the effectiveness of your team, and how strong was the impact, if there was one (direction and strength of impact)?

Because situations may vary substantially during missions, we ask you to consider the most prototypical periods of the mission when assessing the items.

Example: Situational Uncertainty

a. If situational uncertainty was Not Relevant (NR) for your team, then circle NR and proceed to the next question. Otherwise, if it was relevant, please answer the questions associated with this item.
b. Consider the most prototypical period of team functioning and assess the amount of situational uncertainty that accompanied that period of the mission.
c. If the level of situational uncertainty had a very high, positive effect on your team, then circle +4 under Direction and Strength of Impact; conversely, if the level of situational uncertainty had a very strong, negative effect on your team, then circle -4 under Direction and Strength of Impact. If there was neither a positive nor a negative impact, then circle 0. Ratings between -4 and +4 should be used to represent intermediate levels of impact.

Command Team Effectiveness Survey

MISSION FRAMEWORK

The mission framework is defined by *situational uncertainty, stress potential, constraints* and *stakes*.

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Situational uncertainty *The lack of information about, or lack of understanding of objects and their properties in the area of responsibility*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Stress potential *stress potential due to operational consequences*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Constraints *External factors that limit the range of the team's actions or autonomy*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Stakes *The immediate and long-term consequences of mission outcome*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TASK

The task parameters considered are *complexity, workload, goal clarity* **and** *goal stability*.

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Task complexity *Having to deal with rapidly evolving situations, multiple and concurrent tasks, Uncertainty*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Workload *Refers to the cognitive / physical demands of the task*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Lack of goal clarity *The extent to which the team does not understand it's goals objectives, and priorities*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Lack of goal stability *The extent to which goals change significantly over time.*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

ORGANISATION

The organisational context parameters considered are *goal congruity, command structure, autonomy,* **and** *organisational support*.

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Congruity of the team's mission and organisational goal *Degree of fit between goals*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Clarity of command structure *e.g. report to national and international chain of command, staff and line versus matrix structure*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Autonomy *Freedom of action allowed by the organisation*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Organisational support *Degree to which the organisation assists and sustains the team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TEAM LEADER

The *Leader's skill, knowledge,* and *personal goals* are considered as they may have an effect on team processes **and** performance.

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Leader skills *Abilities as they relate to the military task at hand and with the leader's abilities as leader of the team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Leader knowledge *Knowledge, wisdom and experience, which the leader brings to bear in accomplishing the mission*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Match of personal goals to organisational goals *The degree to which the leader's goals are not different from the organisational goals*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TEAM MEMBERS

The team members are also defined by their *skills, knowledge,* **and** *personal goals***.**

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Team member skills *Abilities held by individual team members which enable them to complete their tasks within a team setting*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Team member knowledge *Knowledge, wisdom and experience, which most team members bring to bear in accomplishing the mission*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Match of personal goals to organisational goals *The degree to which the team members' goals are similar to the organisational goals*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TEAM

Teams have different dimensions that include: *team composition, team size, team architecture, team maturity,* **and** *team goals***.**

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Appropriateness of team composition Mix of people on the team according to the mission goal	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Appropriateness of team size *Is the team sufficiently staffed to accomplish the task within the given constraints?*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Appropriateness of team architecture *Refers to the structure of the team; the distribution of subtasks and roles, and how they are related to each other*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Team maturity *External Refers to the extent to which team members have worked together and developed as an intact team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Match of team goals to organisational goals *The degree to which the team goals are similar to the organisational goals*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TASK FOCUSED BEHAVIOURS

Task-related processes include both productive and corrective behaviours: *managing information, assessing the situation, making decisions, planning, directing and controlling, monitoring progress,* and *liaising with other command teams.*

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Managing information *The team's way of handling information or knowledge*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Assessing the situation *Includes perceiving, recognizing, and anticipating environmental elements or events*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Making decisions *Includes creating multiple options, choosing among alternatives, and implementing optimal solutions*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Planning *Formulating actions necessary to achieve a goal*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Directing and controlling *Processes that occur between planning and attaining a goal*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Liaising with other command teams *Includes developing and maintaining contact or communication with other command teams(e.g. by giving and receiving information, coordinating activities)*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TEAM-FOCUSED BEHAVIOURS

The team-focused processes include both productive and corrective behaviours: *providing and maintaining vision, maintaining common intent, interacting within the team, motivating, adapting,* and *providing team maintenance.*

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Providing and maintaining vision *Generating and preserving direction and purpose*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Maintaining common intent *Preserving a shared sense of goals, objectives, and action*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Interacting with the team *Communication and co-ordination within the team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Motivating *Influencing the direction, intensity, and persistence of team members' behaviours*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Adapting *Using strategies for changing circumstances*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Providing team maintenance *Team building activities that keep the team together*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

Continued on other side of page.

TASK OUTCOMES

The task outcomes considered are the *intermediate goals* **and** *end goals***.**

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Achievement of the intermediate goals *Decisions being made, plans ready according to standards*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Achievement of the end goals *Decisions being made, plans ready according to standards*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

TEAM OUTCOMES

The team outcomes considered are *mutual trust, morale, cohesion, confidence, shared vision* **and** *mutual respect.*

TO WHAT DEGREE ARE EACH OF THESE ELEMENTS PRESENT IN YOUR COMMAND TEAM?	Rel	Quality/Magnitude	Direction and Strength of Impact
Mutual trust *Team members have trust in each other's competence, loyalty, and dedication to the teamwork*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Morale *The team members are willing to continue their work, even under adverse conditions*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Cohesion *The team members act as a team instead of as individuals, they feel attracted towards the team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Collective confidence in achieving the goal *Team members have a strong belief in the effectiveness of the team*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Shared vision *In general, team members perceive the task that they have to accomplish in the same way*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4
Mutual respect *Despite differences, the team members try to understand each other*	NR	1 2 3 4 5	-4 -3 -2 -1 0 1 2 3 4

This page intentionally left blank.

Annex D Daily Questions for Ardent Sentry 06 Exercise

Instructions

To the question administrator (the following paragraph is not to be presented to the people actually taking the survey):

For each of the four days from May 8th to May 11th we have prepared a short set of questions which are intended to test situation awareness, susceptibility to biases in decision making, the operator's sense that they have an understanding of the commander's intent, and the operations understanding of the role of the other operations centers. Each of these factors has been identified as being important for organizational and individual effectiveness. Each set of questions is intended to be asked after work for that day of the exercise is complete.

May 8th

To the operator:

At the end of each day you will be asked a series of questions about the events which occurred during the exercise. There are two types of questions; one type is factual (e.g. how many soldiers were sent to location x) and the other type is subjective (e.g. how good was your situation awareness). For the factual questions, please respond to the best of your knowledge even though, sometimes, the facts will not be relevant to your role. The goal of the questionnaire is to determine who knows what and we fully expect that people will not have knowledge of certain facts. Many of the questions will require confidence ratings. It is important that you use the confidence ratings in a meaningful way (i.e. please do not check off the same confidence rating for each of your answers). We would like you think of the percentage ratings as a rating of the probability that your answer is correct. That is, if you say you are 80% confident in your answer it would mean that, on average, for all the times you felt that level of confidence in your opinions, you would be correct about 80% of the time. The subjective questions require just a percentage response.

Name:_____ LOCATION_____

1) Which birds have been diagnosed with H5 in Canada?
   a) Canadian Geese
   b) Chickens
   c) Turkeys

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

2) How many turkeys were affected by H5N1 in NE?
Answer:_____ Range (+/- how many turkeys)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

3) What virus do the birds in Quebec have?

a) Unknown
b) H5N1
c) H5N2

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

4) What virus do the turkeys in New England have?

a) Unknown
b) H5N1
c) H5N2

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

5) Which virus is more pathogenic?

a) Unknown
b) H5N1
c) H5N2

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

6) To what extent do you believe that your actions today were consistent with the JTF or Canada COM Commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

7) To what extent do you believe that activities in your Ops Centre today were consistent with the JTF or Canada COM Commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

8) To what extent do you believe that your actions today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

9) To what extent do you believe that activities in your Ops Centre today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

10) As far as your responsibilities are concerned to what extent do you feel you had good situational awareness during today's activities?

0% 20% 30% 40% 60% 80% 100%
(check one)

11) Which scenario was your main focus today? _____
(Please answer the next three questions in the context of this scenario)

12) Currently, which operations centre has the most complete information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

13) Currently, which operations centre has the most up-to-date information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

14) Currently, management of the scenario requires the immediate action from which of the following ops centers?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

15) What is your best estimate of the probability that the H5N1 virus will infect at least one human in Canada within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	

10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

16) What is your best estimate of the probability that the H5N1 virus will infect at least one human in the major urban centre of Toronto within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

17) What is your best estimate of the probability that not even one human in Canada will be infected with the H5N1 within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

May 9th

To the operator:

At the end of each day you will be asked a series of questions about the events which occurred during the exercise. There are two types of questions; one type is factual (e.g. how many soldiers were sent to location x) and the other type is subjective (e.g. how good was your situation awareness). For the factual questions, please respond to the best of your knowledge even though, sometimes, the facts will not be relevant to your role. The goal of the questionnaire is to determine who knows what and we fully expect that people will not have knowledge of certain facts. Many of the questions will require confidence ratings. It is important that you use the confidence ratings in a meaningful way (i.e. please do not check off the same confidence rating for each of your answers). We would like you think of the percentage ratings as a rating of the probability that your answer is correct. That is, if you say you are 80% confident in your answer it would mean that, on average, for all the times you felt that level of confidence in your opinions, you would be correct about 80% of the time. The subjective questions require just a percentage response.

Name:_____ LOCATION_____

1) Who reported that there is a shipment of fake Tamiflu going to the Eastern seaports?

a) CSE
b) RCMP
c) CSIS

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

2) Where is the poultry farm where poultry is dying from an unknown cause?

a) Woodstock
b) Guelph
c) London

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

3) Which virus was confirmed in migratory birds?

a) Unknown
b) H5N1
c) H5N2

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

4) How many backyard flocks are suspected to be infected with the bird flu?
Answer:_____ Range (+/- how many flocks)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

5) To what extent do you believe that your actions today were consistent with the JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

6) To what extent do you believe that activities in your Ops Centre today were consistent with JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

7) To what extent do you believe that your actions today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

8) To what extent do you believe that activities in your Ops Centre today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

9) As far as your responsibilities are concerned to what extent do you feel you had good situational awareness during today's activities?

0% 20% 30% 40% 60% 80% 100%
(check one)

10) Which scenario was your main focus today? _____
(Please answer the next three questions in the context of this scenario)

11) Currently, which operations centre has the most complete information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

12) Currently, which operations centre has the most up-to-date information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

13) Currently, management of the scenario requires the immediate action from which of the following ops centers?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

14) What is your best estimate of the probability that the H5N1 virus will infect at least one human in Canada within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

15) What is your best estimate of the probability that the H5N1 virus will infect at least one human in the major urban centre of Toronto within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	

100%	

16) What is your best estimate of the probability that not even one human in Canada will be infected with the H5N1 within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

May 10th

To the operator:

At the end of each day you will be asked a series of questions about the events which occurred during the exercise. There are two types of questions; one type is factual (e.g. how many soldiers were sent to location x) and the other type is subjective (e.g. how good was your situation awareness). For the factual questions, please respond to the best of your knowledge even though, sometimes, the facts will not be relevant to your role. The goal of the questionnaire is to determine who knows what and we fully expect that people will not have knowledge of certain facts. Many of the questions will require confidence ratings. It is important that you use the confidence ratings in a meaningful way (i.e. please do not check off the same confidence rating for each of your answers). We would like you think of the percentage ratings as a rating of the probability that your answer is correct. That is, if you say you are 80% confident in your answer it would mean that, on average, for all the times you felt that level of confidence in your opinions, you would be correct about 80% of the time. The subjective questions require just a percentage response.

Name:_____ LOCATION_____

1) Which agency corroborated the terrorist claim of responsibility for the RDDs in Windsor?

a) CSIS
b) RCMP
c) CSE

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

2) How many people were likely killed immediately by the RDDs in Windsor?
Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

3) How many people in total will die as a result of the RDDs?

Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

4) Who received a phone call claiming responsibility for the RDDs in Windsor?
a) CVNN

b) Local police
c) RCMP

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

5) Where have fake Tamiflu pills been shipped?
a) St-John
b) Halifax
c) Fredericton

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

6) Fabric Air flight 510's initial scheduled destination was:
a) Washington
b) Baltimore
c) Atlanta

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

7) Two members of RCMP INSET are on route to Windsor from where?
a) Toronto
b) Hamilton
c) London
ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

8) Who received a phone call by a woman warning that her husband and sons were going to hijack a plane?
a) Media
b) Local police
c) RCMP

ANSWER:_____

0% 20% 30% 40% 60% 80% 100%
(check one)

9) To what extent do you believe that your actions today were consistent with the JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

10) To what extent do you believe that activities in your Ops Centre today were consistent with the JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

11) To what extent do you believe that your actions today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

12) To what extent do you believe that activities in your Ops Centre today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

13) As far as your responsibilities are concerned to what extent do you feel you had good situational awareness during today's activities?

0% 20% 30% 40% 60% 80% 100%
(check one)

14) Which scenario was your main focus today? _____
(Please answer the next three questions in the context of this scenario)

15) Currently, which operations centre has the most complete information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

16) Currently, which operations centre has the most up-to-date information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

17) Currently, management of the scenario requires the immediate action from which of the following ops centers?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

18) Given recent events, what is your best estimate of the probability that another RDD will be exploded in Canada in the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

19) Given recent events, what is your best estimate of the probability that another RDD will be exploded in the major urban centre of Toronto within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

20) Given recent events, what is your best estimate of the probability that no additional RDDs will be exploded in Canada within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	

1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

May 11th

To the operator:

At the end of each day you will be asked a series of questions about the events which occurred during the exercise. There are two types of questions; one type is factual (e.g. how many soldiers were sent to location x) and the other type is subjective (e.g. how good was your situation awareness). For the factual questions, please respond to the best of your knowledge even though, sometimes, the facts will not be relevant to your role. The goal of the questionnaire is to determine who knows what and we fully expect that people will not have knowledge of certain facts. Many of the questions will require confidence ratings. It is important that you use the confidence ratings in a meaningful way (i.e. please do not check off the same confidence rating for each of your answers). We would like you think of the percentage ratings as a rating of the probability that your answer is correct. That is, if you say you are 80% confident in your answer it would mean that, on average, for all the times you felt that level of confidence in your opinions, you would be correct about 80% of the time. The subjective questions require just a percentage response.

Name:_____ LOCATION_____

1) Where is the flight carrying 50% passengers with flu-like symptoms from?
   a) New York
   b) Boston
   c) Bangor

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

2) How many passengers are on board the flight carrying the sick passengers?
Answer:_____ Range (+/- how many passengers)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

3) The Chickens in Woodstock, NB died from what?
   a) H5N1
   b) H5N2
   c) Unknown

ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

4) A flight landing in what city has passengers sick enough to go to the hospital?
   a) Fredericton

b) Halifax
c) Montreal
ANSWER:_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

5) How many people were likely killed immediately by the RDDs in Windsor?
Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

6) On May 9th we asked you to estimate the number of people likely killed immediately by the RDDs in Windsor. What was your estimate then?

Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

7) How many people in total will die as a result of the RDDs?

Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

8) On May 9th we asked you to estimate the number of people likely to die as a result of RDDs. What was your estimate then?

Answer:_____ Range (+/- how many people)_____

Confidence in answer

0% 20% 30% 40% 60% 80% 100%
(check one)

9) To what extent do you believe that your actions today were consistent with the JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

10) To what extent do you believe that activities in your Ops Centre today were consistent with the JTF or Canada COM Commander's intent commander's intent?

0% 20% 30% 40% 60% 80% 100%
(check one)

11) To what extent do you believe that your actions today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

12) To what extent do you believe that activities in your Ops Centre today were consistent with the intent of higher commands?

0% 20% 30% 40% 60% 80% 100%
(check one)

13) As far as your responsibilities are concerned to what extent do you feel you had good situational awareness during today's activities?

0% 20% 30% 40% 60% 80% 100%
(check one)

14) Which scenario was your main focus today? _____
(Please answer the next three questions in the context of this scenario)

15) Currently, which operations centre has the most complete information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

16) Currently, which operations centre has the most up-to-date information on the scenario?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

17) Currently, management of the scenario requires the immediate action from which of the following ops centres?

a) Canada Command
b) JTFC
c) JTFA
d) GOC
e) Other_____

18) Given recent events, what is your best estimate of the probability that another RDD will be exploded in Canada in the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

19) Given recent events, what is your best estimate of the probability that another RDD will be exploded in the major urban centre of Toronto within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	
90%	
100%	

20) Given recent events, what is your best estimate of the probability that no additional RDDs will be exploded in Canada within the next week? Check the appropriate block:

1 in one billion	
1 in one million	
1 in one thousand	
1 in one hundred	
10%	
20%	
30%	
40%	
50%	
60%	
70%	
80%	

90%	
100%	

Annex E – FORMEX Primers and Rating Scales

JSTAFF Collaboration & Decision Making Formex -A

Date

Meeting Subject Matter		*(as applicable to OPP - i.e., COA, Op O)*
Meeting Type		*(briefing, etc.)*
Meeting Level		*(strategic, operational, tactical)*

Collaboration Profile

	1	2	3	4
Number of participants				
Time/ Location				
Meeting Duration				

1 = least complex
4 = most complex

Communication Profile

	Oral	Graphic	Paper Text	Spread-sheet/table	Email	Web Portal	Phone	Liaison Officer
Medium/Mode								

Computer Use Profile

	Word Proc.	Graphic	Database	Spreadsheet	Statistical Analysis	Comms SW	Web	Other
Applications								

Complexity Rating

Decision Making

	1	2	3	4
Consequence of error				
Reversibility of the decision				
Adequacy of the information available				
Whether there is a set procedure or decision tree to follow				
Whether there is a body of similar, past decisions to compare to				
The extent to which judgement is required to make an appropriate decision				

Problem Solving

	1	2	3	4
Complexity of assessing the solution				
Complexity of identifying the problem				
Complexity of identifying the solution steps				
Overall complexity of the problem				

Finding Information

	1	2	3	4
The complexity of locating the desired information				
The complexity of extracting/ processing the information				

Risk Analysis

Urgency

	1	2	3	4
Time sensitivity				

Control

	1	2	3	4
Ability of JSTAFF to control resources involved				

Criticality

	1	2	3	4
The impact of meeting on OPP				
The impact on operations/resources				

Duration

	1	2	3	4
Duration for which the issue has on JSTAFF's radar				

Notes (ie., meeting output, chair style, overall tone and flow of meeting, etc)

The objective of the meeting formex was to capture those details that are pertinent to the meeting as a whole. The following meeting details are recorded at the outset:

1. Meeting Subject Matter – describes the overall topic of the meeting as it pertains to the OPP. Examples include development of the Op Order and discussing COA.
2. Meeting Type – describes the type of meeting. Examples include briefings, brainstorming sessions, etc.
3. Meeting Level – pertains to the level of the overall subject matter addressed during the meeting. Three possible levels are strategic, operational, and tactical.

Collaboration Profile – articulates the collaborative aspects of the meeting, including:

1. Number of participants – captures the number organizations/individuals attending the meeting
2. Time/location – describes the meeting within a time-location matrix. Along the time axis, the meeting can be conducted either synchronously (i.e., real-time) or asynchronously (i.e., non-real time). The location axis categorizes the meeting as either being conducted co-located (i.e. face-to-face) or distributed (e.g., via e-mail, teleconference, video conference).
3. Meeting duration – expresses the length of the meeting

Each dimension is captured in accordance with the Characterization Rating Scale:

.

Dimension	Level 1	Level 2	Level 3	Level 4
Number of participants (departments)	JSTAFF only	Other DND depts	One OGD/Coalition	More than one OGD/Coalition
Number of participants (persons in attendance)	0 to 10	10 to 20	20 to 30	30+
Time/Location	Synchronous/ Co-located	Synchronous/ Distributed	Asynchronous/ Co-located	Asynchronous/ Distributed
Duration	uner 15 min	under 1 hour	under 2 hours	more than 2 hours

The information on Complexity Rating for each dimension is presented in table format below:

DECISION MAKING – Complexity Rating Scale

Dimension	Level 1	Level 2	Level 3	Level 4
Consequence of error	Little or no consequence of error.	Errors have some minor consequence, e.g., some loss of money or time, but can be rectified with some minor work plan, inconvenience or cost.	Errors have significant consequences, e.g., significant loss of money or time, but can be rectified.	Errors have significant consequences that are not rectifiable or are only rectifiable at significant cost.
Reversibility of the decision	Decision easily reversed.	Decision can be reversed with some inconvenience or difficulty; decision is reversible but options are reduced.	Decision can be reversed with significant difficulty.	Decision cannot be reversed, or it can be reversed only with major (legal, financial, health) consequences.

Adequacy of the information available	All information relevant to the decision is known.	Most information relevant to the decision is known.	Information about significant elements relevant to the decision is uncertain.	Significant information relevant to the decision is not known.
Whether there is a set procedure or decision tree to follow	There is a set procedure or decision tree to follow, any bases for exceptions are clearly specified.	There is a set procedure or decision tree to follow but there are also grounds for exception that require some discretion or interpretation.	There is a set procedure but it provides significant scope for discretion or interpretation.	There is no set procedure or decision tree.
Whether there is a body of similar, past decisions to compare to	There are similar past decisions that are directly applicable and that are available to the decision maker.	There are similar past decisions but some extrapolation or analysis is required to apply them to the present decision.	There are past decisions but they provide limited guidance only due to their small number or their limited comparability to the present decision.	No comparable past decisions on which to base the present decision
The extent to which judgement is required to make an appropriate decision	Limited or no judgement needed to make an appropriate decision.	Need to consider several well- defined factors to make an appropriate decision in cases where the consequence of error is low. May involve using technical knowledge.	Need to consider many factors in order to make an appropriate decision. These factors may be less well defined and the consequence of error may be higher than at Level 2.	Significant judgement required in making an appropriate decision.

PROBLEM SOLVING – Complexity Rating Scale

Dimension	Level 1	Level 2	Level 3	Level 4
Complexity of assessing the solution	Check that problem has been solved.	Assess efficiency and effectiveness of solution that was used.	Assess efficiency and effectiveness of solution that was used and identify changes needed.	Solver must identify or create criteria for assessing effectiveness of the solution.
Complexity of identifying the problem	All appropriate information is provided to solver.	Procedures are provided for determining the nature of the problem.	Solver must determine what procedures are to be used to identify the nature of the problem.	Solver must create procedures to identify the nature of the problem.
Complexity of identifying the solution steps	Procedures are given for matching a solution to the problem, once it has been identified.	Solver has to determine which of several available solutions are most appropriate.	May have to modify existing procedures for solving the problems to meet new needs.	Solver must create procedures for solving the problem.
Overall complexity of the problem	Limited number of factors.	Broad range of factors, most of which are clearly defined.	Broad range of factors, some of which may be vague or ambiguous.	Unpredictable and contradictory factors play a role.

FINDING INFORMATION – Complexity Rating Scale

Dimension	Level 1	Level 2	Level 3	Level 4
The complexity of locating the desired information	Consulting established sources, e.g., looking up a phone number in a phone book, calling an airline information number for flight schedule information, consulting a manual, calling a software hotline. Source is supplied to worker, e.g., telephone interviewer who is supplied with the names or numbers to call.	No established source but a source can be easily identified, e.g., workers may enquire of their supervisor or co-workers, "Who would know . . .?"	Worker must conduct a more complex search for the information, e.g., locating witnesses to a crime, setting up appropriate interviews for a research project, collecting appropriate samples for environmental tests.	Information from several different sources must be brought together or there is no source; the information must be created, e.g., conducting research to find a new vaccine.
The complexity of extracting/ processing the information	Information is usable in the form in which it is obtained, e.g., a phone number, a flight time, information on which key indents text in a particular word processing package.	Simple processing, such as selecting information according to some predetermined criteria, e.g., putting together a bibliography, making a list of suppliers for some service in a particular area.	Some analysis required. Information must be understood to be acted upon.	Complex analysis or synthesis. Information from various sources is synthesized. Information is used in the process of generating a solution to a problem. Information is created.

RISK ANALYSIS - Complexity Rating Scale

Urgency	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Time sensitivity	Need a solution/decision within month or more - need solution/decision within 3-6 months of mission begin.	Need a solution/decision within a week - need solution/decision before mission begin	Need a solution/decision today - need a solution/decision before next step in OPP	Need a solution/decision during meeting - need a solution/decision to complete this step in OPP.
Criticallity	**Level 1**	**Level 2**	**Level 3**	**Level 4**
The degree of impact to ops/mission/resources?	little impact to ops/misson or resources	impact to ops/misson or resources - could derail or be life threatening to DND	impact to ops/misson or resources - could derail and be life threatening to DND	impact to ops/misson or resources - could derail and be life threatening to DND and civilians
The degree of impact on OPP	Little impact on OPP	Possible impact on current phase of OPP	Possible impact on more than one phase of OPP	Possible impact on total OPP
Control	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Ability of JSTAFF to control resources	JSTAFF involved	Other DND involved	OGD or Allies involved	NGOs/Allies/OGD/host govt
Duration	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Duration for which the issue has on JSTAFF's radar	today/now	within the week	within past month	more than one month

FORMEX B – Formex for Topic Discussed

JSTAFF Collaboration & Decision Making Formex - B

Date

Meeting Subject Matter *(as applicable to OPP - i.e., COA, Op O)*

Discussion Topic

Discussion Level *(strategic, operational, tactical)*

Product Profile

	Source of Data	Oral	Graphic	Database	Paper Text	Statistical Analysis	Spread-sheet/table	Email	Web Portal	Phone	Liaison Officer
Input											
Input											
Input											
Output											
Output											
Output											

Interaction Profile	Lead/ Initiate	Take Message	Seek New Information	Seek Clarification	Discuss Information	Provide Information	Corroborate Information	Refute Information	Instill Under-standing	Co-ordinate Work	Facilitate	Negotiate
J1 - Personnel												
J2 - Intelligence												
J3 - Operations												
J4 - Logistics												
J5 - Specialists												
J6 - Information Management												
J7 - Doctrine, Training, LL												
J8 - Finance												
J9 - CIMIC, Evironmental H&S												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
Other DND _____												
OGDs _____												
OGDs _____												
OGDs _____												

The letter S or O in cells above denote subjective vs objective information presentation/request a check denotes interaction but unable to identify S or O.

Risk Analysis

Urgency	1	2	3	4
Time sensitivity				

Control	1	2	3	4
Ability of JSTAFF to control resources involved				

Criticality	1	2	3	4
Impact of issue on OPP				
Impact on mission/resources				

Duration	1	2	3	4
Duration for which the issue has on JSTAFF's radar				

Notes

The objective of the topic formex was to capture those details that are pertinent to the individual topics being discussed by meeting participants. The following topic details are recorded at the outset:

1. Meeting Subject Matter – describes the overall topic of the meeting as it pertains to the OPP. Examples include development of the Op Order and discussing COA.
2. Discussion Topic – describes the individual topic of discussion
3. Discussion Level – pertains to the level of the subject matter addressed during the meeting. Three possible levels are strategic, operational, and tactical.
4. Date

Product Profile – details the characteristics of the input and output products supporting the topic of discussion. This includes capturing the source of the data product as well as its medium for distribution.

Interaction Profile – presented in a table that matches individuals with the purposes for oral communication. Purposes for oral communication include the following:

1. Lead/Initiate – direct the discussion surrounding the current topic
2. Take message – capture messages and relay essential information, by phone or in person, to other workers
3. Seek new information – pose questions or make a request to obtain additional information to address a gap in knowledge
4. Seek clarification – request an amplification or explanation to address a particular issue
5. Discuss information – exchange of information and/or opinions between meeting participants
6. Provide information – supplying new information that has been requested by another meeting participant
7. Corroborate information – confirm or substantiate existing information
8. Refute information – dispute or contest existing information
9. Instill understanding – instruct
10. Co-ordinate work – organize and synchronize work being (or to be) conducted by multiple meeting participants
11. Facilitate – guide and maintain focus for a group conversation and/or collaboration
12. Negotiate – resolve conflict. Does not refer only to formal negotiations.

Within the Interaction Profile matrix, each interaction instance will be captured with either an "S" to denote subjective information or "O" for objective information. Differentiating between the two types of incoming information will help to analyze two styles of decision making:

1. Intuitive decision making – based on subjective incoming information; and
2. Analytical decision making – based on objective incoming information

It was the intention to capture complexity and risk for each topic discussed, however, it became apparent through the data collection process that this would not be possible. The meeting topics were either few and thus complexity correlated to that of the meeting or the meeting topics were many with a short attention given to them and thus the full degree of complexity was not able to be captured.

Annex F Questionnaire used to evaluate COPlans

In order to determine if the questionnaire results are dependant of user's OPP knowledge or computer familiarity, we will ask the following questions:

(All) How would you rate your level of knowledge of the OPP? (Evaluation Scale)

1) None
2) Weak
3) Good
4) Very Good
5) Excellent

(All) How would you rate your level of experience to execute the OPP? (Evaluation Scale)

1) None
2) Weak
3) Good
4) Very Good
5) Excellent

(All) Name the computer-based systems other than Microsoft Office that you are using on a regular basis (Descriptive)

The following questions have been prepared for the three groups (Potential User Group, COPlanS Group, Control Group) to confirm the experiment hypothesis mentioned at the beginning of the document:

Tempo and Synchronization:

(All) How would you rate the tempo of the decision making process of your group? (Evaluation Scale)

1) None
2) Weak
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what degree did COPlanS effect the tempo of the decision making process? (Evaluation Scale)

1) Significantly detract
2) Somewhat detract
3) Neutral
4) Slight Improvement
5) Significant Improvement

Information Sharing:

(All) How would you rate the sharing of information amongst your team members? (Evaluation Scale)

1) Poor
2) Weak
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what degree did COPlanS improve the sharing of information amongst your team members? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improved
5) Significantly Improved

Staff Assessment:

(All) How would you rate the quality of the JOPG staff assessment (mission analysis)? (Evaluation Scale)

1) Poor
2) Satisfactory
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what level did COPlanS improve the quality of the JOPG staff assessment (mission analysis)? (Evaluation Scale)

1) Significantly detract
2) Slightly detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) Was the mission analysis brief data obtained with COPlanS complete and accurate?

1) Incomplete and Inaccurate
2) Complete and Inaccurate
3) Incomplete and Accurate
4) Complete and Accurate

(For COPlanS users) Did COPlanS support the collaborative mission analysis activity effectively? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) To what level did the use of COPlanS effect the mission analysis task workload? (Evaluation Scale)

1) It reduced the work load
2) It made no difference
3) It increased the work load

COA/Decision:

(All) How would you rate the quality of the COA produced by your team? (Evaluation Scale)

1) Poor
2) Satisfactory
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what level did COPlanS improve the quality of the COA produced by your team?

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(All) How would you rate the quality of the COA comparison that was done by your team? (Evaluation Scale)

1) Poor
2) Satisfactory
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what degree did COPlanS support the COA comparison? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) To what level did COPlanS help structure the process for deciding on a COA? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) To what level did COPlanS support the production of the decision briefing? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) Was the decision brief created in COPlanS complete and accurate?

1) Incomplete and inaccurate
2) Complete and inaccurate
3) Incomplete and accurate
4) Complete and accurate

General:

(All) How would you rate the quality of the decision making process of your group? (Evaluation Scale)

1) Poor
2) Satisfactory
3) Good
4) Very Good
5) Excellent

(For COPlanS users) To what level did COPlanS improve the quality of the decision making process? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) To what level did COPlanS support the operational planning process? (Evaluation Scale)

1) Significantly Detract
2) Slightly Detract
3) Neutral
4) Slightly Improve
5) Significantly Improve

(For COPlanS users) How would you rate the effort required to learn how to use COPlanS? (Evaluation Scale)

1) Too High (too demanding)
2) High
3) Appropriate
4) Easy
5) Very Easy (Not demanding)

(For COPlanS users) What would be the most appropriate training length needed to be able to use COPlanS? (Evaluation Scale)

1) One day
2) Half a week
3) A week
4) Two weeks
5) More than two weeks

(For COPlanS users) What training method would you recommend for COPlanS? (Descriptive)

(For COPlanS users) How often did you experience problems with COPlanS? (Evaluation Scale)

Never
Sometimes
Frequently

(For COPlanS users) How significant were the problems you encountered with COPlanS? (Evaluation Scale)

1) Minor
2) Moderate
3) Significant

(All) Comments: (Descriptive)

This page intentionally left blank.

Annex G Situation Awareness Rating Technique (SART)

Situation Awareness Rating Technique (SART) is a method for the subjective estimation of SA. In SART, 10 dimensions are evaluated with a 7-point Likert scale (1: low to 7: high). More specifically, these 10 dimensions are regrouped into three distinct categories:

– Understanding of the situation (U): This category includes queries related to the information quality, information quantity, and the level of situation familiarity.

– Demands on attention resources (D): This category includes queries related to the instability of the situation, complexity of the situation, and variability of the situation.

– Supply of attention resources (S): This category includes queries related to the participant's level of arousal, their capacity to focus their attention, their spare mental capacity level and their level of concentration.

Consequently, the SART questionnaire does not only allow the evaluation of the SA level that someone has about a situation but also the cost related to this level.

5. Instability of the situation

How changeable is the situation? Is the situation highly unstable and likely to change suddenly (high), or is it very stable and straightforward (low)?

LOW 1 2 3 4 5 6 7 HIGH

6. Complexity of the situation

How complicated is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)?

LOW 1 2 3 4 5 6 7 HIGH

7. Variability of the situation

How many variables are changing in the situation? Is there a large number of factors varying (high) or are there very few variables changing (low)?

LOW 1 2 3 4 5 6 7 HIGH

8. Arousal

How aroused are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)?

LOW 1 2 3 4 5 6 7 HIGH

9. Concentration of attention

How much are you concentrating on the situation? Are you bringing all your thoughts to bear (high) or is your attention elsewhere (low)?

LOW 1 2 3 4 5 6 7 HIGH

10. Division of attention

How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

LOW 1 2 3 4 5 6 7 HIGH

11. Spare mental capacity

How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (high) or nothing to spare at all (low)?

LOW 1 2 3 4 5 6 7 HIGH

12. Information quantity

How much information have you gained about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)?

LOW 1 2 3 4 5 6 7 HIGH

13. Information quality

How good is the information you have gained about the situation? Is the knowledge communicated very useful (high) or is it a new situation (low)?

LOW 1 2 3 4 5 6 7 HIGH

14. Familiarity with the situation

How familiar are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

LOW 1 2 3 4 5 6 7 HIGH

Annex H Mission Awareness Rating Scale (MARS)

Mission Awareness Rating Scale (MARS) is a situation awareness assessment technique designed specifically for use in the assessment of SA in a military exercise. It comprises two separate sets of questions based upon the three level model of SA (perception, comprehension and projection) [5]. It also includes two subscales, the content subscale and the workload subscale.

– The content subscale consists of three statements designed to elicit ratings based upon ease of identification, understanding and projection of mission critical cues. The fourth statement is designed to assess how aware the participant felt they were during the mission.

– The workload subscale also consists of 4 statements. The first three are designed to assess how difficult, in terms of mental effort, it is for the participants to identify, understand and project the future states of the mission critical cues. The fourth statement is designed to assess how difficult it was mentally for the participant to achieve the appropriate mission goals.

Note that one critic of MARS is that it could be argued that rather than measuring SA itself, MARS is actually rating the difficulty in acquiring and maintaining SA. It is exactly for that reason that the SA concept is measured with different metrics. While SART provided information on the participants' subjective evaluation of their own SA, MARS provided information on their capacity to acquire and maintain SA. The other SA metric used in this experiment, QUASA probe, provided information on the participants' knowledge about critical events in the situation (SA content).

15. Please rate your ability to identify mission-critical cues in this mission.

___ very easy – able to identify all cues

___ fairly easy – could identify most cues

___ somewhat difficult – many cues hard to identify

___ very difficult – had substantial problems identifying most cues

16. How well did you understand what was going on during the mission?

___ very well – fully understood the situation as it unfolded

___ fairly well - understood most aspects of the situation

___ somewhat poorly – had difficulty understanding much of the situation

___ very poorly – the situation did not make sense to me

17. How well could you predict what was about to occur next in the mission?

___ very well – could predict with accuracy what was about to occur

__ _ fairly well – could make accurate predictions most of the time

___ somewhat poor – misunderstood the situation much of the time

___ very poor – unable to predict what was about to occur

18. How aware were you of how to best achieve your goals during this mission?

___ very aware – knew how to achieve goals at all times

___ fairly aware – knew most of the time how to achieve mission goals

___ somewhat unaware – was not aware of how to achieve some goals

___ very unaware – generally unaware of how to achieve goals

The last four questions ask how difficult it was for you to detect and understand important cues present during the mission.

19. How difficult – in terms of mental effort required - was it for you to identify or detect mission-critical cues in the mission?

___ very easy – could identify relevant cues with little effort

___ fairly easy – could identify relevant cues, but some effort required

___ somewhat difficult - some effort was required to identify most cues

___ very difficult – substantial effort required to identify relevant cues

20. How difficult – in terms of mental effort – was it to understand what was going on during the mission?

___ very easy – understood what was going on with little effort

___ fairly easy – understood events with only moderate effort

___ somewhat difficult – hard to comprehend some aspects of situation

___ very difficult – hard to understand most or all aspects of situation

21. How difficult – in terms of mental effort – was it to predict what was about to happen during the mission?

___ very easy – little or no effort needed

___ fairly easy – moderate effort required

___ somewhat difficult – many projections required substantial effort

___ very difficult – substantial effort required on most or all projections

22. How difficult – in terms of mental effort – was it to decide on how to best achieve mission goals during this mission?

___ very easy – little or no effort needed

___ fairly easy – moderate effort required

___ somewhat difficult – substantial effort needed on some decisions

___ very difficult – most or all decisions required substantial effort

Annex I Quantitative Analysis of Situation Awareness (Quasa)

Quantitative Analysis of Situational Awareness (QUASA) is for the assessment of the situational awareness of individuals participating in C2 experiments and exercises. The technique combines both objective queries (true/false probes) and subjective self-ratings of the confidence for each probe response.

– In QUASA, each query results in two specific answers, one related to the participants' knowledge about the situation and the other, to the participants' level of confidence about the knowledge. The pairing of these two answers allow the identification of :

  - Optimal situations where participants know the answer (good answer) and are highly confident about it (very high level of confidence).
  - Situations where participants do not know the answer (answers provided randomly) and are aware of their knowledge lacking (very low level of confidence).
  - Problematic situations where participants do not know the answer (wrong answer) but are highly confident they know (very high level of confidence).

In our DCP, participants were asked to fill out the SART, MARS and QUASA questionnaires at the end of the day. The QUASA questionnaire included different sets of probes (based on injects used during the day). They had to fill out the questionnaires based on the state of mind they were in when executing the C2I2 processes.

The use of three types of SA metrics (SART, MARS and QUASA) allowed the evaluation of situation awareness on three different aspects:

– SART: Provide a subjective evaluation of the level of SA based on the difference between the level of knowledge understanding (U) and the level of attentional resources required and available (D – S);

– MARS: Provide an evaluation about the participants' capacity to develop and maintain an adequate level of SA during the mission execution;

– QUASA: Provide an evaluation of:

  - What the participants surely knows about a situation;
  - What the participants don't know about a situation;
  - What the participants wrongly think they know about a situation.

Results of these questionnaires addressing different aspects of SA were correlated together in order to provide a more complete assessment of this concept.

Day 1 with answers

Statement	True/False	Confidence
Yellowknife suffers mechanical breakdown and is unable to board the RCMP Emergency Response Team.	True False [ ] [x]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Yellowknife suffers mechanical breakdown. YEL cannot perform function, limited to 4 kts.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Combination of illness among civilian police and unanticipated pace of protests and incident response has made ISU seek support from the military.	True False [ ] [x]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
ISU asks for DND assistance to help man the Pemberton police holding cells.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Irate environmentalists attempting to contact DND called JTFP Ops Centre in Esquimalt to make complaint.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Irate environmentalists accuse soldiers of shooting a deer in Cypress Provincial Park.	True False [ ] [x]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Vancouver Sun reporter calls JTFG to ask about the incident in Cypress Provincial Park.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
A suspicious package, looks like a bomb, has been discovered on the routine 1615 Sea Bus departure from Lonsdale Quay.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Van Comd has directed that the ferry, which has 103 passengers on board, dock immediately at Horse Bay.	True False [ ] [x]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
Van Comd has the situation under control and the capacity to evacuate the ferry.	True False [ ] [x]	Very low to Very high [ ] [ ] [ ] [ ] [ ]
About the bomb threat incident, the suspicious package turns out to be a fake bomb with no explosive material, but clearly meant to look like a bomb.	True False [x] [ ]	Very low to Very high [ ] [ ] [ ] [ ] [ ]

Annex J Communication & Collaboration Questionnaire

Generally speaking, how would you rate the communication activities between team members?

very difficult	1	2	3	4	5	**very easy**

Generally speaking, the timeliness level of the information exchanged among team members was:

very bad	1	2	3	4	5	**very good**

Generally speaking, the clarity level of the information exchanged among team members was:

very bad	1	2	3	4	5	**very good**

Generally speaking, the correctness level of the information exchanged among team members was:

very bad	1	2	3	4	5	**very good**

Generally speaking, the completeness level of the information exchanged among team members was:

very bad	1	2	3	4	5	**very good**

Generally speaking, other team members knew who was doing what within the team:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, other team members knew what I was doing during the task execution:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, I was waiting for others' inputs to do my job:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, other team members were waiting for my inputs to do their jobs:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, jobs executed by different team members were complementary within the overall team execution:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, jobs executed by different team members were supportive:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

Generally speaking, team members were exchanging feedbacks:

Never (during this day)	1	2	3	4	5	**All the time (during this day)**

This page intentionally left blank.

Annex K Evaluation of the C2 applications

Please enter the position occupied during the mission: _____

Please enter your level of experience (in years) related to this position or a comparable one: _____

About the Command View application:

Please enter your level of experience with the application:

Novice 1 2 3 4 5 Expert

Please enter your percentage of use compared with the two other applications: _____

Most of the time, the clarity of the information provided by this application was:

Very Bad 1 2 3 4 5 Very Good

Most of the time, the timeliness of the information provided by this application was:

Very Bad 1 2 3 4 5 Very Good

Most of the time, the correctness of the information provided by this application was:

Very Bad 1 2 3 4 5 Very Good

Most of the time, the completeness of the information provided by this application was:

Very Bad 1 2 3 4 5 Very Good

This page intentionally left blank.

List of acronyms

AS06	Ardent Sentry 2006
C2	Command and Control
C4ISR	Command, Control, Computers, Communications, Intelligence, Surveillance and Reconnaissance
C2I2	Command and Control Intelligence and Information
CCSE	Common Command Support Element
CF	Canadian Forces
CFC	Canadian Forces College
CFEC	Canadian Forces Experimentation Centre
CFOPP	Canadian Forces Operational Planning Process
COA	Course of Actions
Comd-Net WG	Command Network Working Group
CONOPS	Concept of Operations
CPX	Command Post Exercise
CP	Command Post
CTEF	Command Team Effectiveness
DCP	Data Collection Plan
DND	Department of National Defence
DomOps	Domestic Operations
DRDC	Defence Research & Development Canada
DSTO	Defence Science & Technology Organization
DWDM	Dense Wavelength Division Multiplexing
EX PG	Exercise Pegasus Guardian
GUIDEx	Guide for Understanding and Implementing Defence Experimentation
ICC	Integrated Command Centre
IC2CE	Integrated Command and Control Collaborative Environment
ISU	Integrated Security Unit
JCDS 21 TD	Joint Command Decision Support for the 21st Century Technology Demonstration
JCDS 21 EXP 1	Joint Command Decision Support for the 21st Century Experiment 1
JIIFC	Joint Information Intelligence Fusion Capability

JIMP	Joint Inter-Agency Multi-National Public
JTFG HQ	Joint Task Forces Game Head Quarter
JTFP/G	Joint Task Forces Pacific / Games
LAN	Local Area Network
MARS	Mission Awareness Rating Scale
MECCS	Major Events Coordinated Security Solutions
MSEL	Main Scenario Event List
NATO	North Atlantic Treaty Organization
NDCC	National Department Command and Control
NORAD	North American Aerospace Defence
NORAD USNORTHCOM	North American Aerospace Defence US Northern Command
OGD	Other Government Department
OR	Operational Research
PSTP	Public Security Technical Program
QUASA	Quantitative Assessment of Situation Awareness
RCMP	Royal Canadian Mounted Police
RFA	Request for Assistance
SA	Situation Awareness
SART	Situation Awareness Rating Technique
SME	Subject Matter Experts
SOP	Standard Operating Procedure
S&T	Science & Technology
ToA	Transfer of Authority
TTCP	The Technical Cooperation Program
V2010	Vancouver 2010

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence R&D Canada – Valcartier
2459 Pie-XI Blvd North
Quebec (Quebec)
G3J 1X5 Canada

2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED
NON-CONTROLLED GOODS
DMC A
GCEC REVIEW: JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Joint Command Decision Support for the 21st Century Technology Demonstration Project: Lessons Learned from Experiment, Demonstration and Training Events (U)

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Breton, R.

5. DATE OF PUBLICATION (Month and year of publication of document.)

October 2011

6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)

150

6b. NO. OF REFS (Total cited in document.)

17

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – Valcartier
2459 Pie-XI Blvd North
Quebec (Quebec)
G3J 1X5 Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

15at

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DRDC Valcartier TR 2010-226

10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) Over the course of the Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD) project, two exercises, one experiment and two demonstration sessions have been executed. This document presents the lessons learned from these five events. From these lessons learned, five guidelines have been identified. These guidelines should support the development of measurement protocols for military environments and appropriate metrics based on the constraints related to the environment in which the measurement should occur and the requirements related to each metric application.

(U) Durant le projet Joint Command Decision Support for the 21st Century Technology Demonstration (JCDS 21 TD), deux exercices, une expérience et deux sessions de démonstration ont été exécutés. L'objectif de ce document est de présenter les leçons apprises au cours de ces événements. À partir de ces leçons, cinq lignes directrices ont été définies afin de supporter le développement de protocoles de mesures pour les environnements militaires et l'identification de mesures appropriées basée sur les contraintes environnementales et les besoins particuliers de chaque mesure concernant son application.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Command and Control; Decision Support Systems; Human Performance; Measures of Performance; Experiment; Training; Prototyping; Joint Environment

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
De science et de technologie pour
la défense et la sécurité nationale

www.drdc-rddc.gc.ca