DRDC Toronto No. CR-2004-053

Towards a Trial Plan for Evaluating the COMDAT TD: Final Report

by:
M. L. Matthews and A. R. Keeble

Human*systems*® Incorporated
111 Farquhar St., 2nd floor
Guelph, ON N1H 3N4

Project Manager:
Kim Iwasa-Madge, P.Eng., CHFP
(519) 836 5911

PWGSC Contract No. W7711-017703/001/TOR
Call-Up 7703-10

On behalf of
DEPARTMENT OF NATIONAL DEFENCE

as represented by
Defence Research and Development Canada - Toronto
1133 Sheppard Avenue West
North York, Ontario, Canada
M3M 3B9

DRDC Toronto Scientific Authority
Sharon McFadden
(416) 635-2189

31 March 2004

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **31 MAR 2004**	2. REPORT TYPE	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Towards a Trial Plan for Evaluating the COMDAT TD: Final Report**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence R&D Canada -Ottawa,3701 Carling Ave,Ottawa Ontario,CA,K1A 0Z4**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report outlines a series of evaluations and analyses of the COMDAT TD with a view to conducting future trials to evaluate its potential impact upon operator performance in the Operations Room of the Halifax Class frigate. Specific issues commented upon include the operator-machine interface, the logistics of integrating the TD into a suitable trial environment, the availability of existing scenario elements to provide a suitable evaluation context, the types of performance measures that could be feasibly implemented and options for the format and location of future trials. Le présent rapport décrit une série d’évaluations et d’analyses portant sur la démonstration de technologie d’aide aux décisions de commandement (COMDAT) en vue de mener d’autres essais pour évaluer les répercussions susceptibles d’influencer le rendement des opérateurs qui travaillent dans la salle des opérations des frégates de la classe Halifax. Les questions commentées ci-dessous portent plus particulièrement sur l’interface opérateur-machine (IOM), sur la logistique permettant l’intégration de la démonstration de technologie dans un environnement d’essai approprié, sur la disponibilité d’éléments de scénario fournissant un contexte d’évaluation pertinent, sur le type de mesure du rendement qu’il serait possible de mettre en place ainsi que sur les options relatives à la structure et à l’emplacement des futurs essais.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**36**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Abstract

This report outlines a series of evaluations and analyses of the COMDAT TD with a view to conducting future trials to evaluate its potential impact upon operator performance in the Operations Room of the Halifax Class frigate. Specific issues commented upon include the operator-machine interface, the logistics of integrating the TD into a suitable trial environment, the availability of existing scenario elements to provide a suitable evaluation context, the types of performance measures that could be feasibly implemented and options for the format and location of future trials.

Résumé

Le présent rapport décrit une série d'évaluations et d'analyses portant sur la démonstration de technologie d'aide aux décisions de commandement (COMDAT) en vue de mener d'autres essais pour évaluer les répercussions susceptibles d'influencer le rendement des opérateurs qui travaillent dans la salle des opérations des frégates de la classe Halifax. Les questions commentées ci-dessous portent plus particulièrement sur l'interface opérateur-machine (IOM), sur la logistique permettant l'intégration de la démonstration de technologie dans un environnement d'essai approprié, sur la disponibilité d'éléments de scénario fournissant un contexte d'évaluation pertinent, sur le type de mesure du rendement qu'il serait possible de mettre en place ainsi que sur les options relatives à la structure et à l'emplacement des futurs essais.

Executive Summary

The COMDAT TD represents a potentially important technology to support the processes of picture-building, contact detection, recognition and identification and track management in the Halifax Class Operations Room. In order to evaluate the potential effectiveness of the TD, it has been proposed that trials be conducted involving navy operators performing mission relevant tasks. This report provides some recommendations on the nature of such trials and how they may be implemented in the future.

The report covers the following major areas:

- An outline of aspects of system usability, functionality and operational performance to be assessed in planned sea and land-based trials;
- Comments upon the existing TD functionality and its implications for an evaluation trial;
- Assessment of the availability of existing scenarios to be used in future trials to evaluate the TD;
- Provision of an overall trial plan for evaluating the TD;
- Identification of requirements for data capture and Operator Machine Interface (OMI) enhancements in order to conduct a future evaluation trial;
- Recommendations for the next steps to be taken in the overall trial plan.

The proposed trial plan comprises three components, each designed to evaluate some aspect of the TD. The first component is a "proof of capability" trial to establish the limits and capabilities of the TD and to provide an operating envelop within which to assess the specific functionality. The second component is a trial that involves Navy subject matter experts (SMEs) performing a cognitive walkthrough of the system functionality to assess its utility and to determine a concept of operations. The third component is a human-in-the-loop trial in which the system functionality is assessed in real time by Navy SMEs using the TD to respond to simulated contact events. Options for conducting such a trial in terms of logistics, scenario events and scope are presented.

Sommaire

La démonstration de technologie COMDAT constitue une technologie qui pourrait se révéler importante pour appuyer les processus d'établissement d'une vue d'ensemble, de détection des contacts, de reconnaissance et d'identification ainsi que de gestion des pistes dans la salle des opérations des navires de la classe Halifax. Dans le but d'évaluer l'efficacité potentielle de la démonstration de technologie, il a été proposé de procéder à des essais auxquels participeront des opérateurs de la Marine accomplissant des tâches adaptées à la mission. Le présent rapport propose des recommandations relativement à la nature de tels essais et à leur application dans le futur.

Voici les principaux éléments du rapport :

- un aperçu des aspects qui seront évalués lors des essais planifiés, en mer et sur terre, relativement à la facilité d'emploi, à la fonctionnalité et à la performance opérationnelle du système;
- des commentaires au sujet de la fonctionnalité actuelle de la démonstration de technologie et de ses répercussions sur un essai pratique;
- un recensement des scénarios actuels, qui serviront dans les essais futurs, à évaluer la démonstration de technologie;
- l'élaboration d'un plan d'essai global pour évaluer la démonstration de technologie;
- l'identification des besoins concernant l'amélioration de la saisie des données et de l'IOM dans le but de mener un prochain essai pratique;
- des recommandations au sujet des prochaines étapes à suivre dans le plan d'essai global.

Le plan d'essai proposé comprend trois éléments, chacun étant conçu pour évaluer certains aspects de la démonstration de technologie. Le premier élément, un essai servant à « prouver la capacité », vise à déterminer les limites et les capacités de la démonstration de technologie et à fournir des paramètres opérationnels à l'intérieur desquels évaluer une fonctionnalité précise. Le deuxième élément met à contribution des experts en la matière (EM) de la Marine qui procèdent à une revue cognitive de la fonctionnalité du système pour en évaluer l'utilité et établir le concept des opérations. Le troisième élément consiste en un essai comportant un chaînon humain et dans lequel les EM de la Marine évaluent en temps réel la fonctionnalité du système en utilisant la démonstration de technologie pour réagir à des contacts simulés. Le rapport présente finalement les options disponibles pour effectuer un tel essai en termes de logistique, d'éléments de scénario et de portée.

Table of Contents

1. Background

Human*systems*® Incorporated (HSI®) has been contracted by DRDC-Toronto to provide advice on future Human in the Loop (HIL) trials to evaluate the Command Decision Aid Technology (COMDAT) Technology Demonstrator (TD). The specific work items that were required to be addressed, as modified by ongoing discussions with the contract authority, were:

1. Review Technical documentation relating to the TD[1]
2. Review TD (at Lockheed Martin Canada (LMC), Montreal) to obtain familiarity and assess requirements for data capture
3. Consult DRDC-Toronto and Atlantic concerning trial requirements and plans for specific threat events to be included in scenarios
4. Determine aspects of system usability, functionality and operational performance to be assessed in planned sea and land-based trials
5. Identify (from previous work) methods and measures suitable for proposed scenarios and data capture capabilities
6. Identify requirements/capabilities for capturing data from TD workstation for human-in-the-loop (HIL) trial
7. Identify desirable changes to TD operator-machine interface (OMI) for HIL trial
8. Review and Assess Suitability of Scenarios developed for Sensor Weapons Controller (SWC) Task Analysis and scenarios used by Navy in Operations Room Team Trainer (ORTT) for use in a future HIL trial. Recommend scenario event requirements.

It should be noted that as the contract has progressed, new information has become available about technical issues concerning the integration of the TD into a test environment. Consequently, the thinking of the Scientific Authority (SA) has evolved concerning the scope and feasibility of future trials to assess the TD and, in addition, practical concerns about timelines and technical logistics have served to constrain the trial options. The thrust of the present document will therefore be to set out a practical trial plan that satisfices[2] the requirement for an evaluation of the TD from an operator perspective and to make recommendations on what aspects can be practically conducted in the remaining COMDAT Cycle III time frame.

---

[1] These work item numbers do not necessarily correspond with the original numbers on the SOW, which has evolved considerably over the course of the contract.

[2] This term was introduced by Herbert A. Simon in his "Models of Man" 1957. It means to obtain an outcome that is good enough. Satisficing action can be contrasted with maximising action, which seeks the biggest, or with optimising action, which seeks the best.

2. Specific objectives

Rather than reporting on the individual work items in the order listed in the previous section, we believe that the material can be best organized under the following headings to better address the overall goals of the SA for this work.

- Outline aspects of system usability, functionality and operational performance to be assessed in planned sea and land-based trials
- Comment upon the existing TD functionality and its implications for an evaluation trial
- Assess the availability of existing scenarios to be used in future trials to evaluate the TD
- Provide an overall trial plan for evaluating the TD
- Identify requirements for data capture and OMI enhancements in order to conduct a future evaluation trial
- Recommend next steps to be taken in the overall trial plan.

3. Aspects of the TD to be evaluated

In previous work, (Matthews, Webb and McCann, 1997) HSI® has recommended that C2 support systems be evaluated along three different, but complimentary, dimensions:

- System usability
- Functional utility
- Operational performance

System usability incorporates issues such as the design of the operator-machine interface (OMI) and usability of system features by the target population. Of central concern is the degree to which both system and interface design match fundamental physical, perceptual and cognitive characteristics of the user. The ISO 9241-11 (1998) standard defines usability as ”the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”. Specifically, effectiveness refers to “the accuracy and completeness with which users achieve goals”, efficiency refers to ”the resources expended in relation to the accuracy and completeness with which users achieve goals”, while satisfaction is ”the comfort and acceptability of use”. Evidently, usability comprises both user performance (i.e. effectiveness, efficiency) and preference (i.e. satisfaction) factors. This definition would also seem to incorporate aspects of functional utility.

Functional utility: concerns the usefulness of the core system functions to the operator in conducting the tasks that the system is designed to support. It is not just the range of functions that is important, but the way they map onto the user’s goals and the task domain.

High system utility and usability are essential for optimal performance of both the system and the user.

Operational performance means the ability of the system to successfully function in the actual operational context: this means that, in the present case, it contributes directly to improved detection, recognition, tracking and identification of contacts. Such improvements may be in terms of higher accuracy and shorter latencies to identify contacts, fewer track management errors, fewer communications and errors, and reduced operator workload. Also considered as part of operational performance is the potential impact of the system on current operational procedures and concepts of operation.

Sometimes included in the evaluation of operational performance is the issue of *user acceptance*, for which Adelman (1992) outlined four important elements:

- Ease of understanding
- Perceived utility
- Perceived reliability
- Effect on decision maker’s confidence

Since the TD is a prototype only, the Scientific Authority has suggested that aspects of *system usability* be excluded from consideration. Hence, this memorandum will focus on issues relating to what may be feasibly accomplished by way of evaluation in the other two domains.

3.1 Functional utility

The core functions of the current TD appear to be as follows. These will provide the initial focus for evaluating utility.[3]

1. Dual display of wide area picture (GCCS-M) and local picture on separate but contiguous monitors. (Currently configured as the Naval Tactical Display –NTD).
2. An MSDF (Multi Source Data Fusion) application that consists of MSDF processes that fuse track reports or processes input data from the following HALIFAX Class information sources.

   Ownship data sources:

- SG-150 radar data
- SPS-49 radar data
- IFF interrogator data associated with the SG-150 and SPS-49 radars
- CANEWS ESM data
- Ownship Navigation data
- North Crossing data

   Input is also processed from the following **remote data sources**:

   - Link-11 data
   - GCCS data

   The MSDF application fuses track data provided by the above sources to estimate track position and velocity, identify the various targets, and compute a measure of confidence in that Identity (ID).

3. Four modes of presentation of the tactical picture: legacy CCS (Command Control System - no MSDF enhancement), MSDF Local Track mode (ownship AWW sensors integrated); MSDF Global 1 Track mode (MSDF Local plus Link 11), and MSDF Global 2 Track mode (Global 1 plus GCCS-M).

4. Tabular display of both pedigree and association data for each track.

5. The NTD displays generic and specific propositions computed by MSDF. Generic propositions comprise a list of attributes including friend, foe, air or surface, along with their probabilities. Specific propositions comprise a list of up to eight possible platform identifications, along with their probabilities, including other qualifying information such as Country, Language etc. Specific propositions also include the ignorance or residual uncertainty as an additional attribute.

6. A visualisation aid is provided to communicate underlying information used to compute contact identity and position. The operator is provided with a single method to display Positional and Identity Uncertainty that uses a "Bar Chart" graphic. The

[3] These are "operational" functions only; functions designed for what appears to be for test and evaluation purposes of non-HIL components are not listed.

graphic comprises 4 vertical bars for each MSDF track. The first bar represents the staleness or Time Lateness, the second represents uncertainty of a target's allegiance, the third bar represents uncertainty of the target's category/subcategory, and the last bar represents the uncertainty of a target's position. The bars are filled by thirds with increasing uncertainty, that is, the higher the level of the bars the greater uncertainty associated with the track analysis.

7. The degree of spatial uncertainty associated with each contact location can be displayed in the form of an ellipse around the center point of the contact plot.

8. A data amplification read-out area (DARO). This provides essential information concerning a hooked track. This includes: amplification data such as detailed identity information about hooked elements; the belief held by the MSDF application for specific target attributes such as information concerning friend, foe or neutral; the belief values are displayed in descending order so that the attribute with the highest belief always appears on the first row; a scrollable window that presents the MSDF application's identification propositions of the hooked MSDF track; the first line of the scrolled window displays the ignorance or uncertainty left in the MSDF application; the second (and any following) line of the DARO scrolled window contains the belief followed by a proposition that is a list of actual target identifications; upon analyst request, the Runtime Display shall display the MSDF belief associated to the allegiance, country, type, subtype and lethality of the hooked MSDF track in the DARO.

3.2 Operational impact

One general consideration in attempting to estimate the operational impact of the TD concerns the way in which the equipped ship will operate in a wider context, whether it be multi-ship, or a standard Canadian Task Group (TG). More complex and different operational issues would arise, if the MSDF fitted ship were to be sailing in a context of either MSDF or non-MSDF fitted ships. For example in terms of external communications, the volume could be reduced, or could increase if ships need to query the source of some of the MSDF information.

Notwithstanding this particular issue, the following areas would seem to be the most directly affected by the introduction of MSDF technology.

Track management issues

1. *Validating continuity of track data:* This involves confirming that all information associated with a given valid track remains with that track at all times (ie the symbology, force track number, ID and all amplifying information). This may include monitoring the track as it manoeuvres to ensure that the symbology does not track off the valid contact. For air tracks, this process is now performed by the ARRO and monitored by the TrackSup and/or the SWC. A perfectly performing MSDF system could in theory overcome some of the existing radar system shortcomings that influence track quality. In which case this task would be redundant. The monitoring of the accuracy of this process by the AWW team may still be required if the MSDF solutions are not reliable, or not trusted, in critical situations.

2. *Validation of symbology:* this entails monitoring the various symbology to see if it represents a valid contact track. Sometimes false tracks are generated, particularly in poor weather, and these may deteriorate with time and need to be manually dropped. MSDF should eliminate the need for this function to be performed, since such false tracks should never be displayed to the operator in the first place. At present the SWC (or ORO) monitors that symbology validity is being adequately assessed. The ARRO or ASPO are responsible for updates to symbology where track quality deteriorates.
3. *Discrimination and ambiguity resolution of multiple tracks on single contact*. This may be required frequently in a high density tracking environment where several TG members are creating and LINKING out tracks. This task is performed by the Tracksup, ASPO or ARRO, and may require external communications.
4. *Managing MSDF tracks and TG Link tracks:* the MSDF's feature of giving each contact an MSDF-specific track number that is unrelated to the local CCS track number or the force track number can have a potentially significant negative impact to operations. Keeping track numbers straight is a critical track management task, thus the team in an MSDF ship may have to do a constant translation between the track #s they're looking at on their MSDF display and the track #s that are being shared throughout the task group on Link. This creates a potential for error if operators mis-associate a track along the way.

Contact recognition and identification issues:

By recognition we mean the process of specifying the particular platform type of the contact, e.g. MIG 25. By identification, we mean the formal operational procedure that results in the contact being assigned a friendly, suspect, hostile or unknown status, with resulting impact on displayed symbology.

1. *The integration of information from different sources to arrive at an ID:* this task is primarily a team responsibility for each warfare area and would conceivably be highly affected by MSDF. This impact of MSDF should result in fewer communications and reduced workload among the team. However, there would be the potential need to do a new task of performing a validity check of the information and solution of the MSDF auto-ID derived from its information fusion. This again would presumably be a team function.
2. *Routine picture monitoring for new tracks and changes in track status*. This is a core function performed by the team, which is particularly alert to tracks that change characteristics that could change the primary identification. Since, under MSDF, the team is less involved in track creation and monitoring quality and validity quality they could potentially be less aware of such changes or have lowered comprehension of the picture. There may also be a possible increase in workload associated with this function and a need for additional team communications. Conversely, the ability of MSDF to take care of such routine functions should provide the team with more capacity to look for important changes to the tactical situation.
3. *Detection of missile separation from a/c and ID as a missile*. At present anyone in the team can make the required zippo call based on available evidence (separated track is made unknown). It is possible that MSDF could make the ID of the missile tracks

sooner (as opposed to unknown). MSDF could particularly show clear benefits in situations of multi-pronged attack.

Picture building issues

1. *Picture maintenance and situational awareness by the ORO*. As part of his responsibility for watching over all aspects of building and maintaining the Maritime Tactical Picture (MTP), the ORO must on occasion switch between local and wide area view provided by GCCS. The provision of this information in the TD in an adjacent display could potentially enhance the task of picture building and maintenance.

Communication issues

1. Whether the MSDF ship is sailing in a TG that is similarly equipped or not with MSDF will have an influence on the volume and types of communications that are involved in track management. A TG that has mixed MSDF capability may require additional communications in order to resolve issues of track ambiguity and duplication.

These issues are re-examined further in the next section, in a review of existing scenario materials that may contain suitable contact situations that could serve as triggers for the TD in a test and evaluation context.

4. Availability of scenarios to evaluate the TD

The SA required that HSI explore the availability and suitability of existing scenarios or scenario events that could be possibly re-employed for a future HIL trial. Such availability would result in reduced logistical costs for preparing for a trial and would have the additional benefit of the validity of the events being already accepted and tested by the Naval community.

Two sources of scenario data were to be explored – the scenarios used to elicit information as part of the SWC and ASWC Task Analyses and scenarios used in the ORTT for Navy team training purposes. With respect to the former, the question of central interest was whether there were additional scenario elements available, that were relevant to the TD evaluation, and which included operational circumstances beyond those employed in the Cognitive Task Analysis (CTA) on Operations Room Officers (ORO) conducted by HSI.

Accordingly, two tasks were performed. The first was a review of the SWC/ASWC scenario documentation, and the second a visit to the ORTT to interview training personnel in order to evaluate existing scenarios for events of potential interest to the TD evaluation.

4.1 Scenarios used for the SWC/ASWC analysis

Two scenarios were used. One was a multi-threat situation in littoral waters, the second a peacetime counter-drug operation. A review of the multi-threat scenario showed that it included very similar elements used in the CTA scenario, which comprised air, long-range missile, surface, sub-surface and land-based missile threats. The make-up of the task group and available resources was also similar to that of the CTA scenario.

The scenario used for the counter-drug situation included detection and identification of air and surface tracks, primarily of civilian origin and also an interdiction/boarding.

Our conclusion is that the scenario elements contained within these scenarios did not add any new event types or situations, relevant to the COMDAT TD evaluation, that were not already part of the knowledge base derived from the original CTA scenarios.

4.2 Scenarios used in the ORTT for Navy training

HSI was given the opportunity to re-visit the ORTT and to brief senior, training personnel on the COMDAT project and possible evaluation strategies. The goal was to solicit some general discussion on how such an evaluation might be conducted in association with the ORTT (which will be discussed in a later section) and to elicit information on scenario events, to be found in existing ORRT training scenarios, that might provide suitable circumstances for assessing the TD.

The resulting discussions gave rise to the following information, which is found in Table 1, below.

	ORTT Function/ Scenario Element	Expected Current Impact	Data Sources	Potential MSDF Impact	MSDF Mode	Notes Comments/Frequency of events in scenario	Suitable for Baseline trial
	Validating continuity of track data						
1	Fighters manoeuvring close to ship	Lose track & generate new tracks	SG-150	Track continuity	MSDF Local Mode (Canadian Patrol Frigate (CPF) Tracks)	Generally not available. Lose track 3-4% only.	Not in ORTT, but these events could be simulated in the MiniSystem
2	Pairs of aircraft manoeuvring close to ship	Symbology switched among valid contacts	SG-150	Track continuity	MSDF Local Mode (CPF Tracks)	2-3x day	YES
3	MIG25 passes directly over ship at 40000 feet		SG-150	Track continuity	MSDF Local Mode (CPF Tracks)	Causes loss of tracking. 2-3x day	YES
	Validation of symbology						
4	High sea state or heavy rain, ship turning	Generation of false tracks	SG-150	False tracks will not be displayed	MSDF Local Mode (CPF Tracks)	NO not in ORTT-at sea only	Not in ORTT, but these events could be simulated in the MiniSystem

Table 1: Potential scenario events available in ORTT training scenarios

	Discrimination and ambiguity resolution of multiple tracks on single contact						
5	Dual tracking of Link tracks	Manual correlation, voice communication	Link-11	No dual tracking	MSDF Global 1 Track Mode	YES-frequent every missile fire.	YES
6	Validity of newly detected missile tracks	Multiple invalid tracks generated on same contact require manual resolution under time pressure	SG-150 SPS-49 Link	False tracks not displayed	MSDF Local and/or Global 1	YES-frequent every missile fire. Also pop-up contacts May prevent false alarms	YES
	Managing MSDF tracks and TG Link tracks						
7	Initial/Force Track number being inadvertently changed as different participating units take responsibility or release tracks to Link	Allowed to happen, Warfare Commander orders number changed back	Link-11	Unknown	MSDF Global 1 Track Mode	Occurs routinely	YES
	The integration of information from different sources to arrive at an ID						
8	ESM correlation to missile(s)	Manual correlation leading to ID	CANEWS, CCS track	Platform recognition aid	MSDF Local Mode (CPF Tracks)	Routinely	YES
9	ESM correlation to aircraft	Manual correlation leading to ID	CANEWS, CCS track	Platform recognition aid	MSDF Local Mode (CPF Tracks)	Routinely	YES
10	ESM correlation to ship	Manual correlation leading to ID	CANEWS, CCS track	Platform recognition aid	MSDF Local Mode (CPF Tracks)	Routinely	YES

	Routine picture monitoring for new tracks and changes in track status						
11	Contact is identified "Suspect" by another unit	Receive verbal cue from other ship, look at MTP	Link-11, external communica tion circuit	Unknown	MSDF Global 1 Track Mode	Routinely	NO In practice the identification would not be queried-assume the other participating unit is doing its job.
12	Ship receives Link track from (MSDF fitted) consort with identification information from source not held by ownship	Query originating participating unit or may just accept or confirm	Link-11	Query originating participating unit, maybe additional discussion	MSDF Global 1 Track Mode	More of an issue when not all ships are MSDF. Have information locally so no need for confirmation. Unlikely.	NO. See above
	Detection of missile separation from aircraft /ship and ID as a missile						
13	Aircraft/ship launches missile(s)	Air track(s) auto-generated, manually identified	SG-150	Since no automatic identification impact is only on track validity	MSDF Local Mode (CPF Tracks)		NO.

	Picture maintenance and situational awareness by the ORO						
14	Contact is identified, certain attributes are now known or assumed	Mental note of attributes, jot down notes	Tacpac, training, intelligence OPGEN messages, Janes	Attributes from MSDF database automatically displayed	All	Difficult to measure, except though completeness of verbal briefings/reports	YES – but would probably require full scale trial.
15	Receipt of locating information on contact of interest via GCCS-M	Manual plot of information in CCS, manual review of MTP	GCCS-M	Auto correlation with MTP	MSDF Global 2 Track Mode	Surface only – routinely updated.	YES
16	Ship's team inputs amplifying information to GCCS-M on contact of interest	Manual transfer of information from MTP to GCCS-M	GCCS-M	Auto transfer of information. But how to associate global tracks with local tracks.	MSDF Global 2 Track Mode	Some routine examples in scenarios.	NO – this mode not supported in TD.

To summarize, of the sixteen events that could be used to assess different aspects of the TD, ten have the potential to be assessed in the ORTT using existing scenarios, and a further two could be assessed using the capabilities of the mini-CSTC.

Thus, it is clear that, if logistical and technical issues can be overcome, there would be adequate event stimuli available in the ORTT to assess a wide range of the TD functionality and its impact on operational performance.

Further, the scenario events outlined above provide the basis for future activity to analyse existing training records for the purposes of benchmarking operational performance on TD relevant, critical tasks.

5. Implications of the existing TD functionality for an evaluation trial

Towards the end of the contract, HSI personnel were able to spend some additional time in a more thorough review of the TD OMI and its functions. The purpose of this visit was to observe the TD functionality, as currently implemented, to identify any obvious, relatively minor changes that might be made to the OMI to better accommodate a HIL trial, and to identify any changes or additions to the TD architecture to facilitate data collection for a HIL trial. Access to the TD was provided in the MiniSystem from 0800-1200, and from 1400-1700. Two LMC support personnel were available to set up the TD and to run a combination of pre-scripted and original scenarios. Nelson McCoy from DRDC-Atlantic, who has fairly detailed knowledge of the functionality of the TD, was also present to observe and assist HSI. LMC engineers were available to be called in Montreal if needed, but no calls were required.

5.1 TD Configuration

As installed in the mini-CSTC, the COMDAT TD differed in two significant ways from the version observed in Montreal.

- The display used to present the COMDAT TD picture was that of the prototype dual display NTD developed by LMC. This over-under, two screen display allowed for the presentation of the GCCS fed picture on the upper display and the MSDF picture on the lower display, as envisioned by the system developers. The contractors were required to gain some familiarity with the use of the NTD display in order to manipulate the TD.
- The same feed from the TD's fusion engine to the NTD also went to the original TD Scientific Interface (SI) used during the sea trial. This allowed for the comparison of the data displayed on the NTD with the same data displayed on the original SI. This helped identify when confusing or incomplete data on the NTD could be attributed to the NTD itself as opposed to the TD fusion engine.

5.2 Observations Relating to the OMI

5.2.1 The Quick Action Buttons (QABs) are a medium/dark blue in colour. Some QABs such as filters are “highlighted” when they are selected. When highlighted, the QABs are a darker blue, but the difference in button status mode is not readily conspicuous. While not a serious issue for a T&E trial, a simple tweaking of the colour palette would allow an operator unfamiliar with the system to more quickly ascertain the selection within the QAB array.

5.2.2 The naming of filter QABs is not intuitive. It is not evident whether the relevant information is filtered in or out. For an operator unfamiliar with the OMI, the results of selecting the filter must be observed in order to understand if the filter is stopping or allowing the display of data. If there are no tracks that readily fit the category, then the operator may be unsure of the status of the filter.

5.2.3 A number of QABs have no functionality as of yet. These should be suppressed for a trial.

5.2.4 It was not possible to determine the degree of CCS functionality available at the NTD by selecting one of the defined positions. Functionality appeared to be somewhat limited (although this has not been found to be the case with the TD available to the SA). Specifically, with the ORO position loaded, the raw SG 150 radar return could not be displayed because the radar was unavailable[4]. Because of this shortcoming, it could not be determined if any symbology for that matter, that was input and observable on a SSD would be observable on the NTD while in MSDF mode. Further tests were not conducted because this was not the focus of the visit.

5.2.5 In the MSDF mode, track numbers are prefixed with alphanumerics related to the sub-mode selected. While in a MSDF mode, the operator could concurrently select CPF tracks, which would allow the display of all legacy CCS tracks on the NTD. If these legacy tracks were sourced from the SPS 49, SG 150 or Link 11 then they would be overlaid on the MSDF tracks. This gave rise to a visual confusion of the two types of tracks.

It was discovered that within a 20nm range, the legacy and MSDF tracks were displayed with a small separation; this was likely due to the different grid mapping used by each of the two systems. Outside of 20nm the two tracks legacy and MSDF tracks were completely superimposed. This made it impossible to determine which symbology was associated with which track, and to discriminate each track number. This confusion could be somewhat overcome by the operator switching between CCS and MSDF display modes, or the operator could choose to display only the legacy CCS tracks. This would also allow an operator to independently determine the CCS track number relating to any MSDF track. Note that the symbology and data in the CCRO reflecting the Standard ID of the tracks may not be the same in the MSDF and CPF Tracks modes. For example, in one of the scenarios a track generated by a MIG was assessed as hostile by the MSDF, whereas the legacy track indicated an unknown status.[5] This is to be expected because the CCS requires an operator (through QAB action) to manually change the Standard ID of a track while MSDF does this automatically.

---

[4] It was unclear whether this was a NTD issue or a CSTC architecture issue.

[5] The SA has noted that although there is not a one to one relationship between CCS and MSDF tracks, it may be possible to provide a quasi link between the tracks (which could be displayed) by looking up the radar track number for the CCS track’s responsible tracker. This capability may be useful for NTD and SSD equipped Ops Room team members to reference the same track, however it would require the operator to conduct one additional step that is currently not necessary.

5.2.6 There is a QAB that allows the operator to suppress the display of a hooked track, making it disappear from the Tactical Situation Area (TSA). Another QAB allows the operator to display the track again after inputting the track number, but there is no tote showing the tracks that have been suppressed, so the operator has to either remember the track number or jot it down. Since the L and G tracks are unique, a G track cannot be recalled, if an L track has been suppressed. An outstanding question is if the operator suppresses an L track, does the system create a G track using the sources it used to create an L track plus Link. This may not occur because the fusion of the WAP and Link only works with fused local tracks. However, this issue should be clarified in the future, for example would the operator have to call the track LXXXX or G1XXXX, or would it be displayed at all?

5.2.7 In the DARO, when the tabs for Generic and Specific Propositions are selected, the lists of possible propositions are fixed in order. We recommend that the order be changed to place the proposition with the highest probability at the top of the list and the remaining propositions in descending order of probability.

5.2.8 All Generic and Specific Propositions were not always displayed on the NTD, despite being displayed on the original SI. This appeared to be an issue with the NTD as opposed to the TD.

5.2.9 There appeared to be an alignment problem with the probabilities associated with a particular proposition. In the best case, the numeric probability figure was displaced upwards from the horizontal as opposed to being directly in line with its proposition title. Frequently, the probability was beside the incorrect proposition for the Generic Propositions of Neutral, Friend etc, while again it was correct on the original SI. Again, this appeared to be a NTD issue.

5.3 Observations Relating to Data Fusion

Six original mini scenarios were run to observe the outcome of basic fusion processes on tracks. Each of these will be described below together with comments on the TD functionality

5.3.1 A single aircraft track was generated and held on the SG 150 radar. During the course of the run, the underlying radar was turned off for approximately ten seconds, as if flying through a null or descending below the radar coverage horizon briefly, and then being redetected by the SG 150 as it flies along the same course and speed. Initially, with only positional and kinematical information, MSDF correctly assigned highest probability to this being an unknown aircraft. It gave it a 0.8 probability of it being one of approximately eight aircraft in the knowledge library, varying from a MIG to a 747-400. When the radar data was turned off, there was no immediate indication provided to the operator that the SG 150 had lost the original track.

When the SG 150 redetected the aircraft, a **new** MSDF track with a new track number appeared on the NTD. Subsequently, we deliberately altered this track to get some separation between tracks, to see how MSDF would cope with this. For several minutes the two tracks, one representing the actual aircraft and valid track, the other only old symbology, continued on apparently independently of each other. The generic and specific propositions associated with the old symbology did not appear to show the disintegration of the data associated with this invalid track. The Area of Probability (AOP) associated with the old symbology remained very small, almost to the point of being a point source. After several minutes, when the run was about to be terminated, it was observed that the old symbology had significantly repositioned and was tracking very close to but not on top of the true track[6]. The track number did not change and for the next few minutes until the run ended there remained two independent pieces of symbology to represent this one aircraft.

The one piece of useful information was that the SG 150 track number (in the DARO) on both pieces of symbology was the same, indicating that the fusion engine was associating the SG-150 track reports to both of these tracks in some manner. In contrast, when the *legacy CCS* was observed, with no operator intervention, the track quality of the lost track could be observed to be dropping in the Close Control Readout area (CCRO), and the track eventually provided the visual indication of deteriorated quality when it began flashing.

In this specific case, with no operator intervention, MSDF did not improve the tactical picture; in fact, it confused the picture for several minutes by displaying two apparently valid tracks when only one was present. An operator unfamiliar with the scenario would have been led to believe that there were two aircraft.

5.3.2 A MIG 29 track was generated and detected first on the SPS 49 and then by the SG 150; the contact then energised its radar. MSDF performed as expected, with data in the DARO indicating when both the SPS 49 and SG 150 held the aircraft. When the MIG radar was energised, the symbology and readouts in the CCRO immediately changed to hostile. The Specific Propositions changed to a very high probability of this being a MIG 29. As expected, nothing changed on the legacy CCS when the radar was detected by CANEWS. In this case, the MSDF worked in the manner intended. However, there were no data relating to the CANEWS intercept observed in the DARO; this absence of CANEWS data was confirmed in other scenarios.[7]

5.3.3 The third, fourth and fifth scenarios were very similar in that they involved two identical aircraft, flying on exactly the same bearing towards own ship but spaced by 5 miles. Once both tracks were observed on the NTD, a hostile radar was turned on from only one of the aircraft, placing two aircraft on the same bearing as the emission. The first time the trailing aircraft turned on its radar. MSDF reacted exactly as it had in the second scenario above, immediately making the second aircraft hostile (which was incorrect, given the uncertainty of the source of the emission). There was no noted change whatsoever in the status of the first aircraft, and there was no indication that the MSDF had any uncertainty as to the originator of the emission.

Speculating that perhaps there was some programmed rule-based component to the MSDF's assignment of ownership to an EW intercept, the next run had the leading aircraft turn on its radar. Again, MSDF immediately assigned the source to the correct aircraft, ignoring the second. The final run repeated the source of the emission as being the trailing aircraft. This time, MSDF associated the emission with the leading aircraft, which implies that the assignment was random, and that MSDF failed to present the correct assessment, namely, that the emission may have come from either the other aircraft, or a source beyond both aircraft. It would be interesting to see the result if a ship based radar was detected on the bearing of an aircraft, or vice versa.

From this demonstration it can be concluded that the MSDF engine was not providing the appropriate information. It should have indicated that all potential tracks on the EW bearing

---

[6] In fact the track was being treated as fused by the TD, but was arbitrarily being displayed as two tracks because of the grid mapping issue identified in 5.2.5 above.

[7] The SA notes that the CANEWS data does appear to be available when such scenarios are run on the TD at DRDC-Atlantic.

should be made suspect. Also, it should indicate that the EW source might be some other contact, as yet undetected by radar, or not showing in the TSA because of the range selected.

5.3.4 This scenario involved making a GCCS track available for fusion when there was an existing local surface contact in exactly the same position. With only positional information to base its decision on, MSDF associated the two tracks, transferring the attribute data from the GCCS track to the local MSDF track. In this case, the MSDF worked in the manner intended.

5.3.5. In general, the bar graphs related to track quality did not seem to provide useful information. The dynamics of the bars going up and down were difficult to relate to actual track events. The bars were invariably at their highest level, and the level of only one of the bars (that relating to uncertainty of positional information) changed with any noticeable frequency during the scenario runs.

5.4 Conclusions

With respect to the OMI, the observations of potential problem areas were relatively minor, and if no changes were made to the MSDF, the operator would not be significantly impeded in his ability to use the MSDF functionality in a future evaluation trial.

The original mini-scenarios that were run were very basic and represented common operational occurrences. They essentially simulated situations that operators on the legacy CCS would in all likelihood correctly assess and deal with promptly. However, several questions were raised concerning MSDF performance in these scenarios. In the first scenario, why did the MSDF engine take so long to bring the two tracks back together[8] – even though it never seemed to completely associate them? Why did the confidence in the propositions or in the AOP not deteriorate on the old symbology? Why, if MSDF eventually did associate the two tracks, did it take so long? Why did there remain two pieces of symbology instead of MSDF “fusing” the two into one with the original track number? In the scenarios involving EW intercepts, how can MSDF randomly assign ownership of the emission without considering other possibilities, including other contacts currently held or contacts beyond those held? It should be noted that these types of problems would significantly impede an operator in an HIL trial.

5.5 Recommendation

Prior to conducting any trial it will be important to establish what aspects of the TD are functioning appropriately when driven by a range of common scenario events, which the TD is supposed to handle. It is possible that the problems observed could result from (a) core problems with the implementation of the MSDF concepts, (b) software bugs, (c) the most recent software version not being implemented in the MiniSystem and (d) implementation of the TD on the NTD. Details of what would be involved in such a trial are presented in the next section.

Without conducting such a proof of capability trial and addressing areas where the TD performs differently then expected, any COMDAT HIL trial would be compromised by questions and frustrations as operators tried to understand what MSDF was doing to the picture.

---

[8] The assumption that an association was made is based on the fact that for each track, the TD showed the same SG150 track number.

6. Recommendations for a trial plan

The review of the COMDAT MSDF suggests that a multi-step process be used to evaluate the technology. Initially, we will consider all steps as preferred options, although ultimately we shall conclude that only a subset might be realistically implemented in the remaining time available in the TD cycle.

The first step would be to conduct a *Proof of Capability* trial, to check on what aspects of the TD functionality are working in accordance with the current design requirements and what will require workarounds or software fixes for subsequent trials. The second step would involve an SME evaluation of the **functional utility** and some aspects of **operational impact**. The third step would be to conduct some form of **HIL trial** to collect performance, operator feedback and other data on a subset of system functions using a limited scenario in a test facility (such as the MiniSystem or ORTT).

Before addressing these options in more detail, it should be noted that the possibility of conducting a comprehensive **evaluation of the TD in a dockside or sea-based trial** had been discussed in an interim report, as required by the initial statement of work. The purpose of such a trial would be to attempt to validate the TD in the specific operational situations for which it was designed to provide the most benefit. Such situations can be characterized by factors such as: degraded radar quality, ambiguous or seemingly conflicting information from different sensor and other sources, high data rates and multiple threats in close temporal proximity. Some of the situations for which MSDF may be best suited might not be able to be replicated anywhere but at sea (e.g. false contacts due to sea clutter, loss of SG-150 tracking on violently manoeuvring a/c). Given that such a trial is no longer being contemplated by the SA, this option has not been explored any further.

This proposed multi-step approach is suggested for a number of reasons. First, the logistical complexity of mounting a full, HIL, real-time data collection trial is not really necessary to obtain SME data on issues such as functional utility. This can be more readily and efficiently accomplished using "table-top" or walkthrough procedures as indicated below. Also, logistical issues such as access to the TD, support personnel and trial participants would be less complex.

By using an incremental approach we will likely uncover potential ways in which the TD may impact upon each team position, when highly experienced SMEs get an opportunity to see the TD. At present, this impact is based on our best estimate given the existing knowledge base. Such potential discoveries at earlier stages of the evaluation, and incorporating them into the evaluation "scenario" will allow subsequent stages to have higher validity.

The general requirements and considerations for conducting these evaluations are described in subsequent sections.

6.1. Trial 1: Proof of Capability[9]

This trial will involve a systematic test of the TD functionality (with the most recent software) in which simple, but representative, operational events are used to stimulate the TD algorithms, using

[9] It is possible that LMC may have already performed such an exhaustive checkout of the TD at the operator task level, in which case, documentation relating to this should be made available to the project and the trial would therefore not be necessary.

mini scenarios that can be readily and quickly constructed. The events would be selected in discussion with the SA and Navy SMEs to encompass the full range of situations in which the TD could be expected to have an impact. The list already outlined in Table 1 should provide the basis for conducting this task.

The evaluation itself would be framed in terms of the tasks that operators would normally perform with the contact events selected. Thus, the evaluation is aimed at the information provided to an operator on the workstation screen, not the algorithm level. This evaluation need not involve actual operators but could be conducted by SMEs available to the test and evaluation (T&E) team.

The goal of the evaluation would be to establish what aspects of the TD are functioning in accordance with design specifications and what are not. Of those functions that are functioning appropriately, the boundary conditions and limitations of their capability need to be established. In this way, contact event contexts that the TD was never designed to handle would not be included in any evaluation trials. Of those aspects that are not working, there would be a need to establish which are readily correctable prior to future trials involving Naval SMEs, and which aspects would require workarounds in future trials.

It is recommended that the trial be conducted in the CSTC using an appropriate TD workstation that provides the maximum functionality and the minimum limitations. Presumably, the NTD would be the workstation of choice if the full TD functionality can be exercised without compromise.

Support from LMC would need to be provided in terms of personnel who have an intimate knowledge of the TD capabilities and functionality and who could provide the required level of software support to implement and run the various scenario elements. The LMC specialist should have a high level knowledge of the software implementation of the TD to be able to provide information to be able to distinguish those circumstances in which the TD is not functioning as designed, from situations in which it is functioning as designed but has not been previously extended to cover more complex conditions.[10]

The outcome of the capability evaluation would be a data matrix comprising the list of scenario MSDF trigger events, the action of the TD, the success or otherwise of the action, and in case of sub-optimum performance, whether the problem can be solved with a relatively simple software fix, or would require a workaround in subsequent trials.

This trial would then provide the necessary quality assurance that the time of Navy SMEs recruited to participate in future trials would not be wasted unduly, and would eliminate the chance that a malfunctioning system would create a negative impression with the strong potential to bias judgments.

6.2 Trial 2: SME Evaluation of functional utility and Concept of Operations

This trial has two major goals. The first goal would be to obtain valid and representative user evaluation of the core system functionality. The second goal is to determine how an operational version of the TD would impact upon existing procedures and Concept of Operations (COO).

[10] For example, the individual provided should understand the parameters of how long it takes the engine to degrade invalid tracks and then associate them with other tracks (example 5.3.1 above) or what logic is applied to assign an EW intercept to a particular contact as opposed to another (examples 5.3.2., 5.3.3).

The approach would be to first conduct a high level review of the major system functions using a "show and tell" walkthrough of all of the system features with Navy SMEs. The next step would have the SMEs interact with the TD and gain some familiarity with the OMI and how each function is implemented and information displayed. To assist in this process, LMC personnel, or others with a high level of familiarity with the system will be required to be present to answer technical questions concerning the functionality. The final step would be to go through a series of realistic operational contact events, lasting possibly a few minutes each, in which the operator imagines how the TD would be used. Data would be collected through the use of structured interviews and questionnaires.

Issues to be covered would include:

- The value and utility of the MSDF functions
- Tasks that would be impacted (favourably and unfavourably)[11]
- Impact upon procedures and concept of operations, and impact upon workload and communications
- Which team member(s) would benefit most from having the TD at their workstation.
- Enhancements that would be required to improve operational task performance
- Acceptance and trust

Logistical Requirements

Personnel: Having given some consideration to the options available for access to SMEs, we recommend that instructors from the Canadian Forces Naval Operations School (CFNOS) be selected to participate in the evaluation process. Given that we anticipate that the evaluation could take half to a full day, and that it would be desirable to have highly experienced individuals who can draw on their long familiarization with existing capabilities and procedures, this really only leaves the option of instructors and/or regular Ops room personnel. Because of potential problems with access to the latter and their tight schedules, we recommend the use of training staff. Our experience in the past is that these individuals have flexible schedules, have been willing to participate in other DRDC sponsored R&D activities, have insight into existing problems and have the imagination to consider future directions of C2 systems.

Location: The location of the initial evaluation would be the MiniSystem, which would be readily accessible by CFNOS SMEs. However, it should be noted that this may be in high demand for technician training and LMC programming. Presumably such access would be arranged by whoever is coordinating the trials.

Logistics: LMC would be responsible for the installation and operation of the TD and for having expertise available during the assessment to address technical questions. All of the core TD functionality would need to be operational to the point of allowing its critical features to be demonstrated and explored by the evaluators. The T&E team would provide a list of contact events and other appropriate demonstration vignettes to LMC in advance of the trial. A whiteboard should be made available to allow ideas to be worked through by the participating group.

[11] This would include a validation of the expected operational areas impacted as outlined in section 3.2, and the detailed contact events listed in Table 1.

The HF contractor would be responsible for organizing the trial in collaboration with DRDC-Toronto and Atlantic, conducting the trial, analyzing the data, debriefing participants and presenting results.

6.3 Options for a HIL trial

Prior to considering the specific format of a HIL trial, we believe that it would be useful to examine what alternatives exist for collecting data that would address the fundamental issue of how to assess the potential value of the TD in terms of picture building and contact management in the operations room. While the term "human-in-the-loop" generally implies real time performance measurement, with a user performing relevant tasks in the real or simulated operational environment, other approaches are possible. Such approaches allow for the collection of other forms of data that would also be meaningful for assessing the value of the TD. We raise these options now, since there is a strong possibility that given technical constraints in implementing a true HIL trial in the ORTT, and the time available left in the COMDAT cycle, alternate, viable approaches should be considered. The alternatives are presented below with a brief overview of the kinds of data that would be generated and the logistical requirements to implement a trial.

To review, the measurement options that may be considered are:

- Operational performance – real time data
  - Full team-complete scenario
  - Small team-isolated scenario events
- Subjective ratings of operational performance and factors such as utility, trust, acceptance
- Development of a task process model

Note that these are not exclusive alternatives, and some combination of the different measurement approaches may prove to be the most practical and viable approach.

Each of these options will now be described.

6.3.1 Real-time data: full team-complete scenario

First, let us review the goal of this form of HIL trial. The primary goal would be to assess the operational impact of the TD, by obtaining performance data when the TD is used within an air warfare team to react to contact events. Such data could then be compared with baseline data collected with the existing system for similar contact situations. The optimum configuration would be to insert the TD into the ORTT[12] and to capture real-time data as the team responds to specific scenario events.

Such operational performance data may be considered to be the "gold standard" when it comes to assessing the TD capabilities in realistic contact event situations. Operational performance data would include the following components:

- Time to perform tasks (e.g. recognize, identify, resolve ambiguity)
- Number of communications among the team (to complete task)

[12] The specific team member who would benefit most from having the TD available would have been determined in Trial 2.

- Incomplete, inaccurate tactical picture (number of ambiguous tracks)

A major consideration concerning the value of such a trial is whether any reliable, valid or meaningful data may be collected with an operational interface that is quite different from the existing CCS functionality and which may entail a different concept of operations. Because of the nature of the implementation, it seems unlikely that a hybrid approach using part legacy CCS displays and the new TD displays would represent a rigorous or fair test of the new functionality. Further such an approach would create confusion for the operators and less than optimum performance. Thus, for any meaningful comparison data to be generated, it is important that operators be thoroughly familiar with the TD functionality and that a new, ad hoc concept of operations be developed to ensure that the TD is being used in a realistic and operationally effective manner.

Therefore, we recommend that the trial have the following components:

- Selection of SMEs with high experience and who can be flexible in thinking about new ways in which the TD would be actually used in an operational environment (we suggest instructors and staff from CFNOS).
- A preliminary session in which the SMEs become thoroughly familiar with the TD functionality
- An overview of the COO developed as an outcome of Trial 2
- A training and practice session with multiple, real time events to allow the SMEs to become thoroughly familiar with the application of the new concepts and the TD functionality and to refine procedures
- A data collection session with multiple events for collecting real time performance data
- A debrief session to collect from the SMEs objective assessments of the utility of the functionality, its acceptability and their confidence in its usage.

The data collection component would involve a selection of those functions of the TD that are amenable to an event-based, or mini-scenario format involving possibly just a small sub team from the Operations Room, possibly including an ARRO, Tracksup, CANEWS operator and SWC. The goal will be to collect some operationally realistic performance data using a modified existing Concept of Operations that has been adapted to the new functionality afforded by the TD.

Logistical requirements

Location: Such a trial would need to be conducted in the ORTT. One possibility would be to conduct the assessment in the second operations room, while the primary ship's team is under assessment or training in ORTT1.

Scenario: The prior analysis of existing ORTT scenarios has shown that there are a number of events that would be appropriate for assessing the TD (see Table 1).

Personnel: Trial participants would be the air warfare team, plus possibly the ORO, drawn from the crew not currently under assessment. Support personnel would be provided as part of the normal complement of training staff that are used to run scenarios in the ORTT.

Procedure: The air warfare team would be given some preliminary training with the TD and revised concept of operations, prior to the onset of the trial. If the trial were conducted in parallel with the scenario unfolding in ORTT1[13], the team would only be involved in reacting to those scenario events that had been pre-selected as being appropriate.

Data capture: Given the uncertainty of whether the TD workstation can be integrated into the ORTT data capture network, other forms of data capture should be envisaged. These include data logging of selected keystrokes and QABs within the workstation itself, video capture of the TD screen and network communications capture either using the existing ORTT capability, or by having the team wear microphones to allow ancillary audio recording. The video capture could either be achieved by attaching a digital graphics capture system to the workstation monitor, or by using a digital video camera aimed at the screen.

TD enhancements for data capture: If feasible software should be developed to capture and time-stamp QAB selections, specific alphanumeric keys and tracks hooked.

OMI enhancements: The majority of these will be dependent upon the available functionality of the TD and specific manner in which it will be used, as identified in Trials 1 and 2. Recommendations based upon a preliminary review of the OMI were outlined in section 5.2. In general, it is recommended that as many of these items as possible be addressed prior to the implementation of Trial 2. Further, any OMI issues arising out of Trial 1 may also need to be corrected, if feasible.

6.3.2 Real-time data: small team- isolated scenario events

As configured in the MiniSystem, the TD allows for the presentation of air contact events of various types. These are generally presented in isolation and do not involve complex scenarios that extend over lengthy periods of time. This capability could be useful in allowing a scaled-down evaluation trial to be conducted. The general format of the trial would be to prepare a series of air contact situations that are expected to provide the circumstances where the TD should have some impact upon operator or team performance. These mini scenario events could then be presented under one of two conditions (assuming the NTD is used as the operator workstation)– legacy CCS and MSDF TD.

Personnel: Logistical support would be provided by CSTC and LMC personnel. In terms of participants and depending upon available Navy resources, the trial could be conducted either with two separate groups of participants for each condition (preferred option), or the same participants could be used, but the scenario events would be different, but comparable for the two conditions.

The specific air warfare team members that would participate in such a trial would be dependent upon the outcome of Trial 2. It is possible that just an ARRO (or TrackSup) and a SWC would be required.

Data collection: Since the mini-CSTC is not instrumented to allow for real-time data collection, we propose that the activities around the workstation be captured on high-resolution video tape[14]. A multi-camera, multiplexed video approach would allow three aspects of the context to be captured: the TD video area, the keyboard/trackball and the general configuration involving the personnel.

---

[13] We assume this to be the case, because we do not believe that there is a capability to run different scenarios in each of simulated operations rooms.

[14] The desirable TD data capture enhancements outlined in the previous section would also apply to a trial conducted in the mini-CSTC. The approach proposed is to show that a data capture trial could still be successfully run, even if such a capability were not available.

Individual team members would also be fitted with a microphone and their communications integrated onto the video record. The record thus captured would then be subjected to post-scenario analysis to extract the required MOPs.

6.3.3 Subjective ratings

In the event that it proves totally impractical to collect real-time, operational performance data in the manner described above, then an alternate approach would be to collect user opinions and ratings on the TD's potential capabilities in aiding air warfare tasks.

The approach would involve creating a series of mini-scenario events of the type outlined in Table 1 and have individuals do a walkthrough of the typical processes that would be involved using the legacy system. They would then have a chance to familiarize themselves with the TD and then walk through those same processes but using the TD functionality.

Questionnaires and rating scales would be employed to address specific points of comparison between the legacy and TD approaches; these would probably include issues specific to operational performance such as: perceived utility, workload, communication load, the tactical picture and estimated task durations. Also addressed would be macro issues such as trust and acceptability. In general, the goal would be to produce *comparative* ratings for the TD to assess it against existing capabilities.

In this case, a reasonable effort must be made within the time and resources available to ensure appropriate validity and reliability of questionnaires and other measurement metrics. Thus, some pre-testing of the proposed metrics will need to be conducted on a representative sample of Navy SMEs.

This approach could also be used to supplement the previous trial approaches in order to obtain a more comprehensive evaluation.

6.3.4 Task process model

In addition to, in compliment with, or instead of the previous approaches, the TD may also be evaluated in terms of a process model by comparing it to the existing, standard operational process. For example, an information flow/decision-action diagram can be created for the detect-to-identify or detect-to-recognise cycle that is common to many contact situations where information is compiled and analysed by the team. This model would comprise the individual tasks and sub-tasks[15] that are required to complete the cycle, their sequencing and interrelationship and their timing. Such a process model could readily be built from analysis of existing ORTT training records and compiled from several instances (with different teams) of specific scenario events to ensure validity and accuracy. Gross performance measures associated with the model can be calculated in terms of the total number of sub-tasks required, the time taken to complete the cycle, the number of communications and the number of personnel involved. Secondary aspects of the model could also be considered in terms of the rated workload associated with each sub-task – thereby allowing for the calculation of average process workload and workload peaks.

The normative process model created in this way could then be used as a baseline for comparing the processes adopted for similar contact events when the TD is used to perform the task.

[15] Decomposed down only to a sufficient level to allow the appropriate, gross and critical task comparison between the legacy and TD systems.

Logistics: For the highest degree of fidelity, the trial would be conducted in the mini-CSTC. If this were not available, the trial could be implemented in any environment that would allow a static, non real-time demonstration of the TD. This could even be done with a Power Point presentation of the TD functionality and screen etc. The concept of operations for using the TD would be introduced and discussed. Once familiar with the concepts, trial participants would be presented with the sort of contact events that have been described in Table 1. They would then be asked to describe the processes that would be used by the team if they had the TD available to facilitate the reaction to the specific contact events.

In general, this approach has the lowest logistical overhead of any of the evaluation methods.

Personnel: Two or three experienced SMEs, either from operational or CFNOS staff, representative of each team position would be required.

6.4 Recommendations

Based upon expected technical constraints for implementing the TD into the ORTT and the timeline available, we recommend a hybrid, approach that combines the mini-CSTC real time performance data collection and subjective assessment together with a secondary analysis of the underlying task process, that is a combination of the options listed under 6.3.2, 6.3.3, 6.3.4.

7. References

Adelman, L. (1992). Evaluating Decision Support and Expert Systems. John Wiley & Sons, Inc: New York.

International Organization for Standardization (1998). ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.

Matthews, M.L., Webb, R.D.G. & McCann, C. (1997). A framework for evaluation of military command and control systems. DCIEM No. 97-CR-24,*Defence Research and Development Canada No. W7711-5-7253.*

Simon H.A. (1957). Models of Man: Social and Rational. John Wiley & Sons, Inc: New York.

8. List of Acronyms

ARRO	Air Raid Reporting Operator (formerly RT1)
ASPO	Anti-Submarine Plotting Operator (formerly RT2)
ASWC	Assistant Sensor Weapons Controller
AWW	Above Water Warfare
CCS	Command Control System
COMDAT TD	Command Decision Aid Technology: Technology Demonstrator
COO	Concept of Operations
CPF	Canadian Patrol Frigate
CSTC	Combat Systems Training Centre
CTA	Cognitive Task Analysis
DARO	Data Amplification Readout Area
HF	Human Factors
HIL	Human in the Loop
HSI	Human*systems*® Incorporated
ID	Identification
LMC	Lockheed Martin Canada
MSDF	Multi Source Data Fusion
MTP	Maritime Tactical Picture
NCOT	Naval Combat Operator Trainer
NTD	Navy Tactical Display
OMI	Operator Machine Interface
ORO	Operations Room Officer
ORTT	Operations Room Team Trainer
QAB	Quick Action Button
SA	Scientific Authority
SI	Scientific Interface
SME	Subject Matter Expert
SWC	Sensor Weapons Controller
T&E	Test and Evaluation
TG	Task Group
TrackSup	Track Supervisor
TSA	Tactical Situation Area

DOCUMENT CONTROL DATA SHEET

1a. PERFORMING AGENCY	2. SECURITY CLASSIFICATION
Humansystems Incorporated, 111 Farquhar St., 2nd floor, Guelph, ON, N1H 3N4 CANADA	UNCLASSIFIED - Unlimited distribution -
1b. PUBLISHING AGENCY DRDC Toronto	

3. TITLE

(U) Towards a Trial Plan for Evaluating the COMDAT TD: Final Report

4. AUTHORS

M. L. Matthews and A. R. Keeble

5. DATE OF PUBLICATION	6. NO. OF PAGES
March 31, 2004	36

7. DESCRIPTIVE NOTES

8. SPONSORSHIP/MONITORING/CONTRACTING/ASKING AGENCY

Sponsoring Agency:

Monitoring Agency:

Contracting Agency: DRDC Toronto

Tasking Agency:

9. ORIGINATOR'S DOCUMENT NO. Contract Report CR 2004-053	10. CONTRACT, GRANT AND/OR PROJECT NO. PWGSC Contract No. W7711-017703/001/TOR, Call-Up 7703-10	11. OTHER DOCUMENT NOS.
12. DOCUMENT RELEASABILITY Unlimited announcement		
13. DOCUMENT ANNOUNCEMENT Unlimited announcement		

14. ABSTRACT

U) This report outlines a series of evaluations and analyses of the COMDAT TD with a view to conducting future trials to evaluate its potential impact upon operator performance in the Operations Room of the Halifax Class frigate. Specific issues commented upon include the operator-machine interface, the logistics of integrating the TD into a suitable trial environment, the availability of existing scenario elements to provide a suitable evaluation context, the types of performance measures that could be feasibly implemented and options for the format and location of future trials.

Résumé

(U) Le présent rapport décrit une série d'évaluations et d'analyses portant sur la démonstration de technologie d'aide aux décisions de commandement (COMDAT) en vue de mener d'autres essais pour évaluer les répercussions susceptibles d'influencer le rendement des opérateurs qui travaillent dans la salle des opérations des frégates de la classe Halifax. Les questions commentées ci-dessous portent plus particulièrement sur l'interface opérateur-machine (IOM), sur la logistique permettant l'intégration de la démonstration de technologie dans un environnement d'essai approprié, sur la disponibilité d'éléments de scénario fournissant un contexte d'évaluation pertinent, sur le type de mesure du rendement qu'il serait possible de mettre en place ainsi que sur les options relatives à la structure et à l'emplacement des futurs essais.

15. KEYWORDS, DESCRIPTORS OR IDENTIFIERS

(U) COMmand Decision Aiding Technology; Operator Machine Interface; Human-in-the-Loop trials; Technology Demonstrator