Title: **Ad hoc Organization of Distributed Picture Compilation and Support for Situation Awareness in Network Based Defence – An Exploratory Experiment**

Authors: **Hilde Hafnor**

Scientist, Norwegian Defence Research Establishment (FFI),

P.O. Box 25, NO-2027 Kjeller, Norway.

Phone: +47 63 80 70 00, fax: +47 63 80 74 49,

e-mail: hilde.hafnor@ffi.no (point of contact)

Runar Normark

Scientist, Norwegian Defence Research Establishment (FFI),

P.O. Box 25, NO-2027 Kjeller, Norway.

Phone: +47 63 80 70 00, fax: +47 63 80 74 49,

e-mail: runar.olafsen@ffi.no

Ad hoc Organization of Distributed Picture Compilation and Support for Situation Awareness in Network Based Defence – An Exploratory Experiment

Hilde Hafnor and Runar Normark

FFI (Norwegian Defence Research Establishment)
P.O. Box 25, 2027 Kjeller, Norway
Phone: +47 63 80 70 00
E-mail:
{hilde.hafnor | runar.olafsen}@ffi.no

Abstract

Network Based Defence (NBD), adapted from the US' concept of Network Centric Warfare (NCW), is currently one of the primary strategic directions for the Norwegian Armed Forces. NBD have accented the importance of having good situation awareness, not only on the individual level, but on the team level as well. The foundation of shared situation awareness is increased access to and sharing of information in the information infrastructure. Our assertion is that the concept of ad hoc organization of information flow is an important capability in this context. This paper reports the results from an exploratory experiment conducted during the NATO exercise Battle Griffin in February/March 2005. The experiment was a part of the Norwegian Armed Forces Concept Development & Experimentation (CDE) program. The experiment presented here explored ad hoc organization of information flow applied to the distributed compilation of a common operational picture (COP). An additional aim was to test new ways of collaboration (peer-to-peer horizontal collaboration) between military units on the tactical command and control (C2)-level. The main discussion in this paper focuses on how new technology and new ways of collaboration affects the situation awareness of decision makers, both at individual and team level.

1. Introduction

The work described here has been conducted as part of a project carried out at FFI (Norwegian Defence Research Establishment). The project is working in the areas of architecture, middleware, data fusion, psychology and organizational informatics to help build better decision-support systems for military commanders in the future Network Based Defence (NBD). Our main objective is to support Norwegian military forces in developing capabilities for ensuring information access and sharing in the information infrastructure to enable better decision-making. The concept of ad hoc organization of information flow is an idea based on a picture compilation concept, developed at FFI (Hansen et al, 2004), that we believe will increase information access and sharing in a more flexible and timely manner than existing systems provide today.

The experiment used a command and control information system (C2IS) demonstrator developed at FFI (Hansen et al., 2003). This demonstrator utilizes Web Services and peer-to-peer technologies, among others. Two new components have been developed for the experiment: A Resource Registry (a type of look-up service providing flexible access to information and services) and a NetViewer (a Graphical User Interface for the Registry) that make resources in the network available for decision makers.

Our main operational idea is that the new technological solutions will increase the ability to establish a Common Operational Picture (COP) in situations where dynamic configuration of forces is necessary. This will increase shared situation awareness. The idea is also that the processes of picture compilation should be tailored to get the most operational value out of the new technological possibilities. Further we aim to explore new ways of collaboration (horizontal collaboration) between military units, given that new technological solutions for this are available. Horizontal collaboration was therefore performed in the area of picture compilation.

The structure of this paper is as follows: Section 2 gives the background and basic lessons learned from previous experiment. Section 3 outlines our approach to studying situation awareness and teamwork. Section 4 gives a short overview of the new technological elements in our Demonstrator. Section 5 describes the experimental methodology (participants and procedure, measures and measurements). Section 6 presents some of the main results. Section 7 gives a discussion and, finally a short summary and conclusions are given in Section 8.

2. Background

Network Based Defence (NBD), adapted from the US' concept of Network Centric Warfare (NCW), is currently one of the primary strategic directions for the Norwegian Armed Forces. The aim of NBD is to *increase* mission effectiveness by networking military entities, enhancing the sharing of information and situation awareness (SA). The foundation of shared situation awareness is *increased access to* and *sharing of* information. In our view, the information infrastructure has an important role as an enabler, organizer and provider of information access and information sharing. Our assertion is that the concept of ad hoc organization of information flow is an important capability in this context (Rasmussen et al., 2004).

The point of departure for the experiment is the assumption that a dynamic model for organizing the information exchange, i.e. ad hoc organization of information flow, is needed to enable network-based operations. This model for information exchange management is based on the pattern of publish and subscribe, allowing dynamic linking of resources in the network. This is a move towards a "post and pull" direction in command and control (i.e. from push to pull-oriented supply chain).

In the spring of 2004 FFI conducted an exploratory experiment in the military exercise Blue Game 2004 concerning ad hoc organization of picture compilation. This effort confirmed (Hafnor & Olafsen, 2004; Rasmussen et al., 2004) that the concept of ad hoc organization is complex, involving organizational (e.g. how things are done) as well as technological aspects. In order to realize the potential of ad hoc organization it is not enough to experiment with technology alone. Especially because the practice of today most likely need to change or at least be adjusted in order to realize the full potential of the ad hoc organization on tactical and operational command and control level. One of our main conclusions was that the human and organizational aspects - in interplay with technology - must be included in our transformation efforts. This required an interdisciplinary research approach. This constitutes the backdrop of the experiment presented in this paper.

3. Situation Awareness and Teamwork: Theoretical Background

Situation Awareness is essential for making good decisions at the right time. The NBD visions have accented the importance of SA, not only on the individual level, but on team levels as well.

Our point of departure for measuring SA in this experiment is Mica Endsley's (1995) model of situation awareness in dynamic decision-making. Endsley's definition is a useful concept that places emphasis on spatial and temporal awareness as a result of attention towards critical aspects of the environment.

Endsley's operational definition of SA includes both cognitive and context variables. That is, SA is conceptualized as a relation between subjective awareness (cognition) and objective situation variables (context). Endsley refers to three levels of SA:

Perception	Level 1:	Perception of relevant elements in the situation
Understanding	Level 2:	Comprehension of the meaning of elements of the situation
	Level 3:	Projection of the status of elements in the immediate future

These levels form a hierarchy with level 1 as the lowest level and level 3 as the highest (expert/most skilled) level.

The definition is mainly addressing individual SA, but also includes aspects of team SA (TSA) in terms of overlap in individual tasks and the sum of individual SA. This forms the main approach in our effort in assessing SA in this experiment. However, SA becomes especially complex when we consider teams (Hauland, 2002), and we wanted to include more process-oriented measures in order to capture some of the dynamics of teams regarding interpersonal relations and team related variables in the situation (what to be aware of, the coordinated distribution of the situation knowledge within and between teams, and so on). In this experiment we based our team definition on Dickinson & McIntyre (1997) teamwork model. This model consists of seven identified teamwork elements and their mutual relations. Three of these elements, i.e. communication, monitoring and coordination, relates to measuring team SA (Hauland, 2002). *Monitoring* in this context relates to observation and awareness of other team members' tasks and performance. *Coordination* refers to the team members adjusting to each other. *Communication* is the component that links the other components. Communication is the link between monitoring the performance of other team members' and providing feedback about that performance. We therefore included a teamwork assessment measure (Paley et al, 2002) regarding teamwork awareness and mutual awareness of tasks performed, in order to capture some of these team related aspects in our analysis.

4. The Technology Demonstrator

One of the most important applications of the information infrastructure is the distributed production of a situation picture. Our C2IS-demonstrator has been developed over years, and supports production of a distributed situation picture. It consists of autonomous Picture Compilation Nodes (PCN) and supports peer-to-peer collaboration, made possible by an experimental mixture of technologies (Web Services and peer-to-peer technologies, among others). The PCNs are accessed through a Graphical User Interface (GUI), i.e. GeoViewer (see figure 3).

In NBD the possibility for the user to dynamically select information sources will be critical. Information services will be published in the information infrastructure, ready to be utilized by anyone interested and authorized. To enable dynamic user-selected information flow based on service lookup, we needed registry functionality that would allow the dynamic announcement and lookup of services in the information infrastructure. Therefore, two new components were developed for the Battle Griffin 2005 experiment: A Resource Registry (a type of look-up service providing flexible access to information and services) and a NetViewer (GUI for the Registry) that makes resources in the network available for decision makers. This effort gave increased functionality for user interaction, and enabled user collaboration within and between the contexts of a PCN. This also allowed the introduction of unstructured information as supplements to ordinary “COP-tracks”.

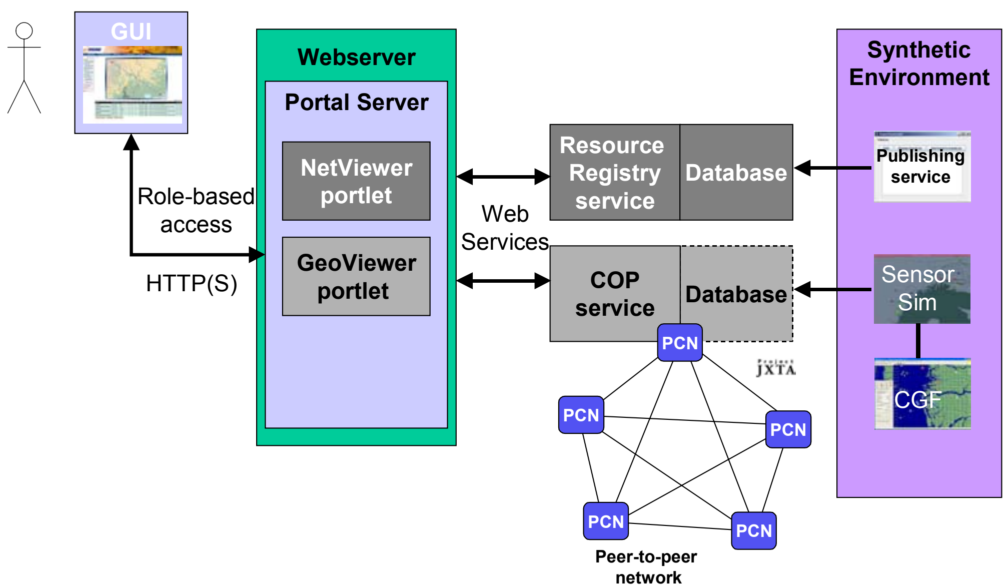

Figure 1 Main elements of the Demonstrator used in the experiment

A synthetic environment simulating a scenario made for this experiment stimulated the PCNs in the demonstrator with data. The synthetic environment consists of two parts: A publishing service providing unstructured information (documents), and a simulation part providing structured information (track information). The simulation part in its turn consists of two applications: A commercial off the shelf computer generated forces package simulating the movements of the entities, and an FFI made simulator simulating sensors observing these entities (Mevassvik et al, 2001).

5. Methodology of the experiment

In this experiment we explored the effects new technology and new ways of collaboration have on the situation awareness (SA) of decision makers. We had constructed a scenario within the overall Battle Griffin 2005 scenario, focusing on protection of an ethnic minority from attacks by paramilitary forces. The task involved collaboration to build a common operational picture (COP), involving land and sea forces, in a simplified simulated escalating military conflict situation. The aim for the participants was to develop an overall picture, a situation awareness, of the whole operational area. The participants collaborated in achieving SA by using the demonstrator.

5.1 Participants and procedure

18 intelligence officers participated and were divided into groups by six. Each group consisted of three teams that collaborated in a decentralized organization (non layered) at tactical level. I.e. the organization designed in this experiment involved no chief in command (no hierarchy). In each team one officer primarily did situation assessment while the other organised the information collection and negotiated on resource allocation. However, there was a floating border between these two functions in each team. The three teams (representing two army units and one coastal ranger unit) were not co-located. The premise for this setting was not only to monitor and collect information for each team's Area of Intelligence Responsibility (AOIR), but also to collaborate in the fusion process of information collected under the collection plan for the whole operational area, thus contributing to a shared situation awareness. At the end the group of three teams were to provide one agreed common picture and an agreed action plan.

All intelligence sources were simulated and delivered information both as structured information (track information that was directly visualized on the screen (in GeoViewer)) and unstructured information (typically humint and other observations) posted on the net and directly available for the user in the NetViewer. All teams were initially given the same information. By linking into the other teams' picture compilation nodes they also shared each others information streams. Also, the NetViewer gave everybody access to all other information (maps, images, historical data, non-structured dynamical data) from sensors (see figure 2).

Chat was the main communication channel between the distributed teams to communicate with each other. Chat was essential to all teams in order to pass information about the understanding of the situation development to each other. In addition the demonstrator supported communication between the teams in allowing team members adding comments to tracks in the situation picture displayed on the screen (see figure 2).

The experiment was run 3 times, one for each group. Thus, we analysed 3 sextets, i.e. a total of 18 subjects (military personnel) by both observations and questionnaires. Each session lasted 4 hours including introduction, on site training and SA measurements. The scenario was played at a speed of four times real time.

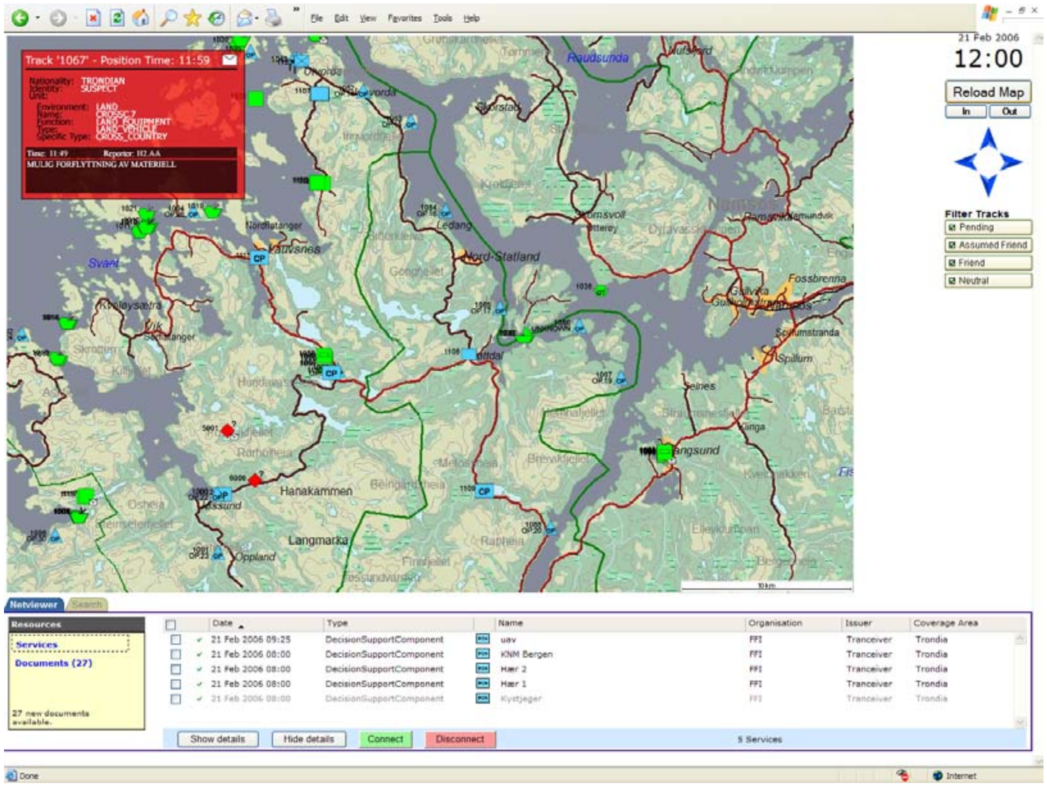

Figure 2 User view: GeoViewer and NetViewer of the Battle Griffin 2005 technology demonstrator

5.3. Measures and measurements

We used a combination of techniques to measure individual SA and team SA. The Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988) covers all three levels of SA. Participants respond to task relevant queries during randomly chosen stops in the simulations. The responses are then compared to the actual state of the environment providing an objective measure of SA. In the present study, the simulation was stopped three times and the participants had five minutes to respond to the queries. We also used SAGAT to measure aspects of team SA and shared SA.

The Situational Awareness Rating Technique (SART) (Taylor, 1990) is a subjective measure of SA. The participants rated their SA on ten different scales. The items are combined into three factors, i.e. Demand, Supply and Understanding, in addition to an overall SA factor that is a composite of the three factors.

- *The Demand factor* concerns the demand on cognitive resources from the context, i.e. the instability, complexity and variability of the situation.
- *The Supply factor* concerns the supply of cognitive recourses, i.e. arousal, concentration of attention, division of attention and spare mental capacity.
- *The Understanding factor* concerns the quality and quantity of information, as well as the degree of familiarity with the situation.

We also supplied our approach to team SA with a teamwork assessment measure regarding teamwork awareness and mutual awareness of tasks performed to capture the dynamics of teams. We used a teamwork assessment measure as a way of assessing the team members' mutual assessment of their teamwork processes. This is also a subjective method that makes the subjects rate the teams in three teamwork behaviours (dimensions)[1] (Paley et al, 2002):

- *Communication*: The ability to provide important information to others (between teams).
- *Backup*: Ability to be aware of each other's workload build-up and react to adjust division of task responsibilities to redistribute workload (within team).
- *Coordination/Information exchange*: The ability to pass critical information to others, thereby enabling them to accomplish their tasks (between teams).

This was also given the subjects after the completion of the scenario run. In addition, during the scenario run, we observed and took notes on team processes.

In this experiment team SA was measured in relation to a simulated military conflict event that required teamwork within and between teams. This was to ensure the relevance of the team SA measures. We also monitored the subjects' ability to use the technology in a collaborative fashion. In addition, the participants evaluated different components of the technology demonstrator by rating to what extent the NetViewer, COP visualisation (GeoViewer) supported their tasks and problem solving activities in the simulation. These measures were then compared to the SA measures.

6. Main results

In this section we present some of the results of our analysis. We used SA measures to evaluate the technology demonstrator, and teamwork behaviour measures to assess some of dynamics of the collaborating teams.

6.1 Individual SA

We used a combination of measures of SA. In large we used SART and SAGAT. Table 1 shows the mean and standard deviation of the SART factors.

Table 1 Mean and standard deviance on the SART factors. Minimum score each factor is 1, maximum is 7 ($N = 18$)

Factor	*Mean*	*S.D.*
Demand	4.1	1.16
Supply	4.5	1.27
Understanding	3.6	1.22
Overall SA (Understanding – (Demand – Supply))	*4.1*	*2.94*

The participants' score on "Overall SA" and the "Demand factor" was average. The scores on

[1] Originally there exist four dimensions in the teamwork assessment measure. In our experiment we omitted the team leadership dimension because of our experimental design of a peer-to-peer organization.

the “Supply factor” was somewhat above average whereas the scores on the“ Understanding factor” was somewhat below average.There was a significant correlation between the “Supply factor” and “Demand factor” in SART ($r = 0.74$, $p < .001$, $N = 18$), but not between any of these factors and the “Understanding factor”. Comparing the groups by using Oneway ANOVA[2] did not reveal any significant differences between the three groups on the different SART factors.

When looking at the single items in SART, the “Arousal item” and “information quality item” deviated most from average. I.e. the participants felt that they were aroused above average (*Mean* = 5.2, *S.D* = 1.06, *N* = 18). Additionally, the quality of the information they acquired was rated as below average (*Mean* = 3.28, *S.D* = 1.60, *N* = 18).

The SAGAT queries concerned all three levels of SA. Table 2 displays the results of selected queries on the different levels. The scores on SA level 2 and SA level 3 queries were higher than the SA level 1 scores.

Table 2 Selected SAGAT queries and responses across participants (18) and stops (3)

	Query	*SA Score (Percent correct)*	*S.D.*
SA level 1			
#1	Select elements or groups of elements and locate them on the map: (*N** = 56)	12%	0.10
SA level 2			
#8	Which of the prioritised information requirements has the highest priority? (*N* = 20)	40%	0.50
#6	What element is the highest threat against the minority civilian population? (*N* = 26)	31%	0.47
#11	Which team is in most need of a UAV resource for information collection? (*N* = 23)	65%	0.49
#15	What is the intention of X forces? (*N* = 10)**	100%	0
#16	What is the intention of Y forces? (*N* = 10)**	62%	0.51
SA level 3			
#10	In which place is it most likely that harassment or attack on civilians will occur? (*N* = 15)**	27%	0.46
#12	What is the most likely course of action (COA) if the x forces should attack the minority civilian population? (*N* = 24)	63%	0.49

* *N* represents the total number of available responses on the particular query across SAGAT stops. *N* varies across the queries because the queries were presented in a random order and the participants had a preset time to respond on the queries.

** Query presented only to participants responsible for analysis (9).

[2] ANalysis Of VAriance (ANOVA).

The participants were in average aware of 12% of the elements in the situation. This is somewhat low but might be a reflection of the complex scenario that included many elements. However it did not prevent them from understanding the intension of the non-compliant forces (100% and 62% correct) and select the correct predicted course of action of these forces (63% correct), although picking the right place of attack was more difficult (27% correct). The standard deviations were high, though.

6.2 Team SA

Team SA is the degree of which team members have the SA to perform their tasks. Team SA is the sum of individual SA. *Shared SA* on the other hand is the degree of which each team member has the consistent understanding of what is going on (Endsley et al., 2003).

We analysed team SA by summing the scores on the SAGAT queries that were common for the participants in each group. Starting with the awareness of the elements (i.e. query #1), the scores differed between the groups. Group A had an average on 40% correct across the stops. The average for group B was 37%. Group C had the lowest average with 25% correct. These results reflect team level 1 SA, i.e. team awareness of elements in the situation. In comparison, the individual level 1 SA scores were lower. The participants in group A had an average of 13% correct (M = 0.13, $S.D.$ = 0.10, N = 18). The figures for group B and group C were 14% (M = 0.14, $S.D.$ = 0.12, N = 18) and 9% (M = 0.09, $S.D.$ = 0.07, N = 18), respectively.

Unfortunately, it was not possible to do any qualified analysis of the team SA on level 2 and level 3. This is due to small number of responses on these queries. We also intended to analyse *shared SA* by comparing the responses on common queries, either within each group as a whole, or between the two participants in each of the three subordinated battle groups. Again there is a problem with the analysis because of small number of responses.

6.2.1 Teamwork assessment and Teamwork Awareness

To achieve high level of mission effectiveness, the members must perform as an effective team. The table below shows teamwork assessment mean scores across team members on teamwork dimension (based on self assessment) within each group. The overall teamwork score is the average of the three dimensions communication, backup behaviour and coordination and information exchange (introduced in chapter 5.3).

Table 5 shows the teamwork assessments scores divided into groups for each run. "Group A" shows the scores for the first run involving teams 1, 2 and 3. "Group B" shows the teamwork score for the second run involving teams 4, 5 and 6. "Group C" shows the teamwork score for the third run involving teams 7, 8 and 9.

The teams differed in their assessment of teamwork behaviour in general ($F\,(2,15) = 8.65$, $p < .01$). Tukey's B Test[3] revealed that the group C differed significantly from the other groups and came out with a significantly lower Teamwork Score.

[3] Tukey's B Test makes pair wise comparisons between groups. We used the test in connection with One-Way ANOVA to se how the groups related to each other.

Table 5 Teamwork assessment score

Teamwork behaviour dimensions	*Group A*	*Group B*	*Group C*
Communication Score	50 %	74 %	29 %
Backup Score	64 %	69 %	48 %
Coordination/Information Exchange Score	60 %	62 %	29 %
Teamwork Score	*58 %*	*68 %*	*35 %*

Teamwork Score Interpretation:

00-19 %	Fair
20-39 %	Moderately Good
40-59 %	Good
60-79 %	Very Good
80-100%	Excellent

An analysis of the specific teamwork dimensions did show significant differences in the "Coordination Score" (F (2,15) = 6.31, p < .01) and the "Communication Score" (F (2,15) = 13.83, p < .001). Again, participants in group C assessed the level of coordination significantly lower than the other groups (Tukey's B Test). All groups differed significantly from each other on the "Communication Score" (Tukey's B Test). Group B scored significantly higher than group A and group A scored higher than group C. There were no significant differences on the "Backup Score".

The results in table 5 support our observation of how the teams interacted with each other during the scenario-runs. While Group B, and to some extent also group A, frequently communicated through chat, Group C was clearly more reserved in their communication and information exchange efforts.

When comparing the team SA level 1 scores and the teamwork score a pattern emerged. Group C differed from group A and B by having a lower teamwork score as well as a lower team SA level 1 score. We could not decide upon the significance of this result. However, we regard this as an interesting observation in our further effort to study collaboration in teams and team SA. We did not find similar patterns between teamwork and team SA level 2 or SA level 3.

6.3 Perceived support of technology demonstrator

The participants found the technology demonstrator easy to use. The mean was above average on a 7-point scale[4] (M = 5.33, SD = 1.50, N = 18) despite that they had received a minimum of training due to limited time.

6.3.1 GeoViewer

Table 6 displays the assessment of GeoViewer. In general, the participants were neutral or responded somewhat positive on all three dimensions of technology support, i.e. that GeoViewer makes problem solving somewhat easier (M = 4.56, $S.D.$ = 1.89, N = 18), efficient (M = 4.61, $S.D.$ = 1.89, N = 18) and effective (M = 4.17, $S.D.$ = 1.92, N = 18) compared to the technology they use today.

[4] All the rating scales concerning technology support are 7-point scales. (1) represent the negative extreme and (7) represent the positive extreme.

Table 6 Appraisals of support by GeoViewer in problem solving activities compared to existing practice ($N = 18$)

Ratings	*Mean*	*S.D.*
Much harder (1) – Much easier (7)	4.56	1.89
Less efficient (1) – More efficient (7)	4.61	1.89
Less effective (1) – More effective (7)	4.17	1.92

The technology support ratings also correlated significantly with many of the SART factors (Table 7).

Table 7 Correlations between SART factors and participant appraisals of support by GeoViewer ($N = 18$)

Support variable	*SART factor*			
	Demand	Supply	Understanding	Overall SA
Ease of use	$r = -.59^{**}$	$r = .68^{**}$	$r = .58^{*}$	$r = .77^{**}$
Efficiency	$r = -.32$	$r = .60^{**}$	$r = .64^{**}$	$r = .65^{**}$
Effectiveness	$r = -.37$	$r = .65^{**}$	$r = .56^{*}$	$r = .66^{**}$

** Correlation is significant at the 0.01 level (2-tailed),
* Correlation is significant at the 0.05 level (2-tailed)

The score on the query concerning SA level 1 in SAGAT did not correlate with the technology support ratings of GeoViewer. Correlations to queries concerning SA level 2 and 3 could not be conducted.

6.3.2 NetViewer

NetViewer makes resources in the network available for decision makers. Table 8 displays the assessment of NetViewer. In general the participants responded in a positive direction on all three dimensions of technology support, i.e. that NetViewer makes problem solving somewhat easier, efficient and effective, compared to the technologies they use today.

Table 8 Appraisals of support by NetViewer in problem solving activities compared to existing practice ($N = 18$)

Ratings	*Mean*	*S.D.*
Much harder (1) – Much easier (7)	4.50	1.89
Less efficient (1) – More efficient (7)	4.67	1.88
Less effective (1) – More effective (7)	4.50	1.79

The technology support ratings also correlated significantly with many of the SART factors (Table 9). However the NetViewer does not correlate significantly with the scores on the Supply and Demand factors.

Table 9 Correlations between SART factors and participants appraisal of support by NetViewer (N = 18)

Support variable	*SART factor* Demand	Supply	Understanding	Overall SA
Ease of use	*r* = -.13	*r* = .37	*r* = .63**	*r* = .47*
Efficiency	*r* = -.17	*r* = .31	*r* = .56*	*r* = .43
Effect	*r* = -.17	*r* = .27	*r* = .76**	*r* = .50*

** Correlation is significant at the 0.01 level (2-tailed),
* Correlation is significant at the 0.05 level (2-tailed)

The score on the query concerning SA level 1 in SAGAT did not correlate with the technology support ratings of NetViewer. Correlations to queries concerning SA level 2 and 3 could not be conducted.

6.3.3 Chat

The participants used chat to communicate with each other. Table 10 displays the participants' assessments of the use of chat. The participants reported that the use of chat did not make problem solving harder or easier compared to the technologies they use today. The same goes with effectiveness. In terms of efficiency, the use of chat may bring a slight improvement.

Table 10 Appraisals of support by chat in problem solving activities compared to existing practice (*N* = 18)

Ratings	*Mean*	*S.D.*
Much harder (1) – Much easier (7)	3.94	1.80
Less efficient (1) – More efficient (7)	4.39	1.88
Less effective (1) – More effective (7)	3.89	1.75

The technology support ratings on chat also correlated significantly with many of the SART factors (Table 11).

Table 11 Correlations between SART factors and participants appraisal of support by chat (*N* = 18)

Support variable	*SART factor* Demand	Supply	Understanding	Overall SA
Ease of use	*r* = -.53*	*r* = .63**	*r* = .39	*r* = .64*
Efficiency	*r* = -.44	*r* = .45	*r* = .49*	*r* = .57*
Effectiveness	*r* = -.60**	*r* = .63**	*r* = .38	*r* = .67**

** Correlation is significant at the 0.01 level (2-tailed),
* Correlation is significant at the 0.05 level (2-tailed)

The score on the query concerning SA level 1 in SAGAT did not correlate with the technology support ratings of chat. Again, correlations to queries concerning SA level 2 and 3 could not be conducted.

8. Discussion

8.1 Individual SA

The purpose of including the SA measures was to evaluate the distributed picture compilation and the support of the technology demonstrator. On the individual SA level the scores on the SART factors were average or close to average. This indicates that the tasks and work were neither too hard nor too easy and they felt that they had the resources to handle the situation. When looking at the individual items in SART, the quality of the information the participants acquired could be better and the situation aroused them clearly above average, which might be a reflection of the intensity and complexity of the situation the participants experienced.

There is a mismatch between the scores on SART and SAGAT. This might be a reflection of the differences between the methodologies. E.g. SART uses self-evaluation reflecting participant confidence and trust in own SA, whereas SAGAT being a more objective, reflecting "actual" SA. In general, the participants' understanding (level 2 SA) and projections (level 3 SA) were better than their awareness of elements in the situation (level 1 SA). In average they were only aware of 12% of the elements. Again, this might be a reflection of the complexity of the situation they experienced, but also of a lacking ability of the technology demonstrator to support the participants in acquiring information and awareness of elements in the situation. Despite the low awareness of elements, the participants were able to understand the situation correctly and choose the right projections to a larger extent.

8.2 Team SA and Teamwork

Our intention in this experiment was to explore how new technology and new ways of collaboration affects the situation awareness, both at individual and team level.

Unfortunately, it was not possible to do any qualified analysis on shared SA. We experienced a problem with the analysis because of small number of responses. This was also very much the case when it came to team SA. It was not possible to do any qualified analysis of the team SA on level 2 and level 3. On level 1 however (team awareness of elements in the situation), we found that team SA seems higher than individual SA.

In general, the teams differed in their assessment of teamwork behaviors. We found that group C differed significantly from the other groups and came out with a significantly lower Teamwork Score. Group C differed from the other groups by performing less coordination and communication activities.

These results are supported by our observation of how the teams interacted with each other during the simulation. Group B had a high frequency of interaction between the teams while group C had almost nothing. This reflects that the self-rating teamwork results varied from "Moderately Good" (group C), "Good" (group A) to "Very Good" (group B). Team SA score level 1 follows the same pattern. Group A and B have a higher level 1 team SA than group C. In this analysis we have not in detail studied the *quality* of the communication (i.e. the substance of what the groups actually communicated). Such a “quality-check” would probably further illuminate these results.

It is worth mentioning that the participants themselves addressed the crucial aspect of “the importance of exercising new work practices”. In addition, they found that using our technology demonstrator and chat as the only communication tool was unfamiliar. In this context, the organization (flat without a chief in command), technology and the way of collaboration differed from what they were used to. This has of course implications on how well the teamwork was performed and to which degree they were able to utilize the new technology to efficiently support the team performance. Adopting new work practices utilizing new technology in a collaborative fashion and having personnel skilled in using it is a long and complex learning process. In our experiment the participants had little time to learn these things. However, the experiment was conducted in a simplified setting in order to highlight certain aspects of importance in the further transformation towards NBD.

8.3 Technology support

In short, the appraisals of the participants were overall positive to the GeoViewer and NetViewer components of the technology demonstrator. These components made it somewhat easier, efficient and effective to perform their tasks and problem solving activities. Positive evaluations of GeoViewer correlated with most of the SART factores. This could be interpreted that the participants who experienced support from GeoViewer also achieved a higher level of SA.

Positive evaluations of NetViewer also correlated with the Understanding factor and overall SA, but not the Supply and Demand factor. The purpose of NetViewer is to provide increased access to resources. Increased access to information would also lead to increased cognitive load and the participants might assess the situation as more complex compared to situations where they have access to less information.

The assessment of chat was more neutral. Chat did not make a difference in terms of making problem solving easier and had no impact on the perceived effectiveness compared to existing systems. However, chat might represent a modest improvement of problem solving efficiency. There were also interesting connections between the assessment of chat and SA. Overall SA in SART correlated positively on all support variables, i.e. good SA is related to positive experiences with chat. More specific, evaluations of chat in terms of ease of use and effectiveness correlated positively with available cognitive resources and negative with demand on contextual demand of cognitive resources. This could mean that the participants that experienced support of chat also had more cognitive resources available and experienced the situation as less demanding compared to participants that did not experience support. In addition, participants that experienced the situation as familiar and experienced that they had the information they needed also perceived the use of chat as more efficient than the technologies they use today. This could be interpreted that when they have relevant experience and relevant information with high quality, the use of chat improves the efficiency. In unfamiliar situations with little information available, chat may not be the technology to use in military tactical and operational command and control.

Unfortunately, these results could not be sufficiently compared to the SAGAT results. However, the SAGAT results revealed that positive evaluations of technology support were not related to awareness of elements in the situation (SA level 1). Unfortunately, we could not analyse the connection between experienced support of technology and level 2 and 3 SA by using SAGAT. Further efforts should focus on this to clarify this issue.

8.4 On methodology

The experiment presented in this paper was explorative. The aim was to provide insight into the relevancy of the ad hoc concept and the use of our technology demonstrator in operative settings. In that sense, the experiment was a success. However there were some methodological aspects that caused problems for us.

First, we had the ambition to study SA including team SA and shared SA. We assessed shared SA to a certain degree but were not able to do a qualified analysis. One cause of this was the lack of responses on SAGAT queries concerning SA level 2 and 3. The respondents had max 5 minutes to respond to the queries. All queries, except for the query concerning the elements in the situation, was presented in a random order to make sure that all queries were covered. Unfortunately, the number of responses was lower than we expected and we could not do all analyses we planned for.

Second, the SAGAT stops are supposed to not to have an impact on the tasks of the participants, at least not to a significant extent. However, we observed that the stops did have a moderate impact on some of the participants and their work. Instead of continuing where they left, many took on other activities, sometime not returning to what they were doing before the stop. Thus, we need to learn more of how the SAGAT stops influences cognitive activities like tactical and operational command and control.

Third, we used a beta software version of SAGAT to design and administer the SAGAT queries, and run into some problems when collecting the data due to errors in the software. The errors will be corrected in the final version of the software.

Forth, the use of the SAGAT methodology does take much effort. To use SAGAT one needs to conduct a goal-directed task analysis (GDTA) for the domain of interest. We based the GDTA on documents, other studies and interviews with intelligence officers. This is time consuming which need to be considered when using the methodology. However, conducting a GDTA is a good way of acquiring and understanding of the work and information requirements in the domain of interest.

For our purpose it might have been enough to focus on the self-assessment measures, e.g. SART, and observation, perhaps in combination with interviews. Performance measures could also be included. On the other hand, the SAGAT methodology may very well serve the purposes of larger experiments with more time and a higher number of participants available, or in studies that compares two or more conditions.

9. Summary and conclusions

An exploratory experiment on ad hoc organization of distributed picture compilation was successfully conducted by FFI during the Battle Griffin 2005 exercise in February/March 2005. The experiment aimed to explore the operational value of selected technological solutions for flexible information sharing in NBD. An operative and technical setting was developed together with a scenario designed especially for this experiment. Essential in the experiment was the use of a C2IS demonstrator developed at FFI. The technology demonstrator supported the participants in their task and problem solving activities.

18 intelligence officers participated and were divided into groups by six, were each group consisted of three teams that collaborated in building a COP. Each group were organized in a flat (peer-to-peer) decentralized organization at tactical level. All elements in the situation were simulated and the experiment scenario was repeated three times, one for each group.

Generally, the participants showed some overconfidence in their individual SA. Due to the complex scenario, the participants were in average only aware of 12 % of the elements in the situation (SA level 1). Despite the low awareness of elements, the participants were able to understand the situation correctly (SA level 2) and select right projections (SA level 3) to a larger extent.

Due to small sample it was difficult to do any qualified analysis on team SA and shared SA. On level 1 however, we found that team SA seemed higher than individual SA. It also seemed that there were variations in the level of teamworking between the groups. However, there was no significant correlation between the level of teamworking and the SA level 1, level 2 and level 3. Yet, when comparing the team SA level 1 and the teamwork, a pattern emerged. We could not decide upon the significance of this result. However, we regard this as an interesting observation in our further effort to study collaboration in teams and team SA. In the experiment the organization, technology and the way of collaboration was different from what they were used to. This had implications on how well the team working was performed and to which degree they were able to utilize the new technology to efficiently support the team performance.

The results show that both the GeoViewer and the NetViewer were appreciated by the participants. These components made it somewhat easier, efficient and effective to perform their tasks and problem solving activities. The experience of support from chat was more neutral. However, chat might represent a small improvement of problem solving efficiency. There were also interesting connections between the assessment of chat and SA. Overall SA in SART correlated positively on all support variables, i.e. good SA is related to positive experiences with GeoViewer, NetViewer and chat.

As an exploratory experiment the experiment has provided few clear answers. However, much has been learned about the possibilities and problems of measuring situation awareness, and several positively interesting observations and questions for further studies have been identified. Yet, these results still support our assumption that the human and organizational aspects - together with technology - must be included in the NBD transformation efforts in order to explore and learn more about this complex interplay.

References

Dickinson T L, McIntyre R M (1997): A Conceptual Framework for Teamwork Measurement. In: Brannick M T, Salas E, Prince C (Eds): Team Performance Assessment and Measurement. Theory, Methods and Applications. Lawrence Erlbaum.

Endsley, M R (1988): *Design and evaluation for situation awareness enhancement*. Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M R (1995): Towards a theory of situation awareness in dynamic systems, *Human Factors* 37, 1, 23-64.

Endsley M R (2003). *Designing for Situation Awareness: An Approach to User-Centered Design*. London & New York: Taylor & Francis.

Hafnor H, Olafsen R (2004): Exploratory experiment of ad hoc organization of picture compilation in NBD: Evaluation of operational value – Blue Game 2004, FFI/NOTAT-2004/01885, Norwegian Defence Research Establishment (FFI) (In Norwegian).

Hansen B J, Mevassvik O M, Bråthen K (2003): *A Demonstrator for Command and Control Information Systems Technology Experimentation* Proceedings of the 8th International Command and Control Research and Technology Symposium, Washington, USA, 2003.

Hansen Bjørn Jervell, Mevassvik Ole Martin, Bråthen Karsten, Rose Kjell (2004): A Picture Compilation Concept for Network Based Defence, FFI/REPORT 2004/00983, Unclassified

Hauland G (2002): Measuring Team Situation Awareness in Training of En Route Air Traffic Control: Process Oriented Measures for Experimental Studies, Ph.D. Thesis, University of Aarhus, Department of Psychology.

Mevassvik O M, Bråthen K and Hansen B J (2001): *A Simulation Tool to Assess Recognized Maritime Picture Production in C2 Systems*, Proc. of the 6th International Command and Control Research and Technology Symposium, Annapolis, USA

Paley M J, Serfaty D, Baker K, Miller P, Baily A, Ganber G, Wan L (2002): Adaptive Performance in Warfighting and Peacekeeping Best Practices Report: Description of the DDD-SASO, Technical Report A003, Army Research Lab, Contract Ng1339-01-0049, Aptima, Inc, Woburn, MA.

Rasmussen R, Gagnes T, Gustavsen R, Hafnor H, Hansen B, Haakseth R, Mevassvik O M, Olafsen R, Rose K (2004): Exploratory Experiment "Ad hoc Organization of Picture Compilation" Conducted during Blue Game 2004: Evaluation Report, FFI/REPORT-2004/01940 (Main Report). Norwegian Defence Research Establishment (FFI).

Taylor, R M (1990): Situational awareness rating technique (SART): The development of a tool for aircrew systems design, Situational awareness in aerospace operations *(AGARD-CP-478)* (pp. 3/1-3/17). Neuvilly Sur Seine, France: NATO – AGARD.