

Utrecht University
Master Applied Cognitive Psychology

THESIS

The Representation of Tactical Information as an Abstraction Hierarchy:
Is the Abstraction Level of Knowledge Representation Related to the Situational Awareness
and the Abstraction Level of Team Communication?

Tim Nederhoff, 3288277

26 November, 2012

First Assessor: Maarten van der Smagt, Utrecht University

Smart Operations Evaluation 2012 project

Under supervision of Wouter Vos and Jan Willem Streefkerk, Institute for Applied Sciences

Netherlands, TNO Human Factors, Soesterberg

Second Assessor: Stella Donker, Utrecht University

Summary

The main goal of a military operational team is to assess the tactical situation in order to act and react to the situation timely and accurately. Over the years, tactical situations have become more complex, including more actors and asymmetric threats. Additionally, military operations have become more comprehensive (i.e., operations require more than just fighting an opponent but also contain activities such as re-constructive work, counterinsurgency, and politics). An important role in achieving situational awareness is extracting knowledge from the system's interface. By applying concepts of *ecological interface design*, environmental information can be presented in a system more effectively. By using an *abstraction hierarchy* in the information representation, a better fit to the operator's mental representation can be realized. This paper describes an experiment where the knowledge representation in a command and control interface is modeled according to the *abstraction hierarchy* by Rasmussen (1985). Measures of situation awareness and communication were analyzed to assess the effect of the implication of abstraction hierarchy in information representation. Results showed no significant difference when hierarchical knowledge representation was compared to conventional knowledge representation. However, trends were directly showing that higher abstraction levels in communication and situation awareness of the tactical situation increases with the application of abstraction hierarchy.

Contents

Contents.....	3
Introduction	4
Background	7
Abstraction Hierarchy.....	8
Situation Awareness	10
Communication	12
Summary.....	14
Interface and Team Communication in Naval Domain	16
TADS Design	16
Team Communication	20
Method	21
Participants	21
Task	21
Manipulation.....	22
Measures	22
Apparatus.....	25
Procedure	27
Analysis	28
Results	30
Discussion	32
References	35
Appendix A. SA summary derived from internal communication	38
Appendix B. Scenario description.....	40
Appendix C. The Four SA Questions and Multiple-Choice Answers	42
Appendix D. Keyword Lists.....	43

Introduction

In naval onboard command and control settings, team members have to plan, direct, control, and monitor the operation they are executing in order to defend the ship and fulfill their mission. These activities have become increasingly complex as the tempo and diversity of operational scenarios increase. Furthermore, technological advances in combat management systems and more and complex data to be processed in time-critical conditions emphasize the high task load of military teams (Grootjen, Neerincx, & Veltman, 2006). Finally, warfare has shifted towards littoral areas which generally contain more commercial air traffic and merchant shipping as opposed to open-ocean areas. Therefore, the tactical situation changes frequently and potential threats can be multiple, with a high degree of uncertainty. That is, asymmetric threats characterized by civilian entities having hostile intentions (De Greef, 2010). A high degree of workload and high levels of uncertainty increase the risk of taking wrong or inappropriate decisions (Wade, 1996). These developments entail risks of misperceiving data from combat management systems and drawing faulty conclusions. As a consequence, the importance of efficient intra- and inter team collaboration has increased. Specifically communication between collaborators has become essential for effective and accurate information transfer between team members. Both the increase in amount of information and the complexity of the tactical situation to be communicated are threatening factors in executing a military operation successfully.

In summary, an operational team has to deal with an increasingly demanding processing task. An efficient and accurate cooperation between a combat management system and the human operator is required for assessing the tactical situation accurately. The team of operators benefits from an efficient, usable representation that provides operators with situational knowledge conform the tactical situation on which the communication is based. Both the content and the structure of how information is represented are important factors for

the perception of the operator. Therefore, such socio-technical systems should be 'ergonomic' to efficiently support the operator in assessing the situation. The interface of such a system could be improved by applying methods derived from the field of *ecological interface design* (see Vicente (2002) for an overview). This design methodology helps operators adapt to change and novelty in the tactical situation by presenting the situational status and the controls of a system in structure that matches the mental representation of the operator.

In this study, we examined the effects of a restructured information representation in a socio-technical system's interface. As a starting point, an interface is designed and based on Rasmussen's work on hierarchical knowledge representation. By representing information in a layered structure (referred to as an *abstraction decomposition space*), a higher compatibility to the operator's mental model is achieved.

Arciszewski and de Greef (2011) developed an abstraction decomposition space that is applied to tactical data in a naval combat management system's interface (referred to as the tactical abstraction decomposition space, TADS for short). We conducted an experiment with experienced Navy personnel in which an implemented TADS interface was compared to a conventional Dutch naval interface. TADS features include higher order information, characterized by abstract indicators for the tactical situation.

With abstraction hierarchy as the manipulation factor, we assessed the abstraction level and the efficiency of verbal communication in a team of operators. In addition, we measured Situational Awareness (SA) that focuses on the awareness of identity, activity, behavior, and mission of military entities (McGuinness, 2004). With the application of TADS, we expect that operators have a better understanding of the tactical situation due to better conformation of knowledge representation in the interface to the operator's mental model. Consequently, we expect that communication increases in efficiency as SA increases. That is,

more information is transferred using communication at a higher density and consisting higher levels of abstraction in its content.

This paper starts with the theoretical background on ecological interface design, situation awareness, and communication in operational teams followed by an introduction to the applied interface by a walkthrough of the interface elements. Manipulation, setup and measures of SA and communication are explained in the method section. Results and implications are discussed in the final chapter.

Background

Starting after the second world war, extensive research on human-machine interaction was conducted (Bennett, Posey, & Shattuck, 2008). Several ‘ergonomic laws’ appear to be fundamental contributors to a cognitive ergonomic design. Stemming from ecological psychology (Vicente & Rasmussen, 1990) and cognitive engineering perspectives (Vicente & Rasmussen, 1992), a theoretical framework for designing human-computer interfaces for socio-technical systems is known as *ecological interface design* (EID, Vicente & Rasmussen, 1990). The term ecological refers to the environment (ecology) that the system is representing. The goal of EID is to provide the operator with tools and information that make the problem and possible solutions perceptually evident to the user. Fundamental in EID, two conceptual methods serve as a theoretical foundation in representing knowledge in a socio-technical system: the *skills, rules, and knowledge based behavior* taxonomy and the *abstraction hierarchy* (described in the next section). From these methods, models can be built for domain specific process control representation (Vicente, 2002).

In information processing tasks, behavior can be defined at three levels: skill, rule, and knowledge based behavior, referred to as the SRK-taxonomy by Rasmussen (1983). With skill based behavior, operators control the system in an automated and low conscious manner. Once the skill is learned, for example changing gear in a manual transmission car, performance is smooth and consumes less cognitive resources. Rule based behavior is characterized by following rules, instructions, and protocols. Knowledge based behavior is the process of reasoning and decision making and is typically related to problem solving tasks. The objective of the SRK-taxonomy is to induce appropriate behavior levels by a system’s interface. This enables the operator to act directly on the interface (skill), mapping work domain constraints to perceptual information in the interface (rule), and providing an

externalized mental model representing the work domain as an abstraction hierarchy (knowledge).

As is demonstrated by research in application domains as process control, aviation, computer network management, software engineering, medicine, and command and control, operators necessarily benefit from ergonomic representation of relevant information (Paradis, Treurniet, & Roy, 1998). For example, an operator in a command and control setting benefits more from problem relevant indicators as opposed to a basic auditory warning signal which only indicates a general system failure. A method of mapping appropriate information to the corresponding task is coined as the *abstraction hierarchy* by Rasmussen (1985) and is discussed in the next section.

Abstraction Hierarchy

When knowledge workers are dealing with a complex technical system (like the operators mentioned in the introduction chapter), they build a mental representation of the system's status in a constructive manner. Rasmussen (1985) found that humans construct their mental models at different levels of abstraction, corresponding to the functioning of the system. These levels can be conceptualized as layers ordered by abstraction, referred to as the *abstraction hierarchy*. Five levels of knowledge representation have been defined: the *functional purpose* level which describes the overall goals and purposes of the entire system, the *abstract functions* level, describing the laws and principles governing the system or situation, the *generalized functions* level concerning the plans, the way how the overall goals are achieved, the *physical functions* level describing the physical world in terms of capabilities and roles of objects, and the *physical form* level which describes the most elementary information of a system or situation. For example, Figure 1 shows an abstraction hierarchy mapped to military command and control information (derived from Arciszewski and De Greef (2011)).

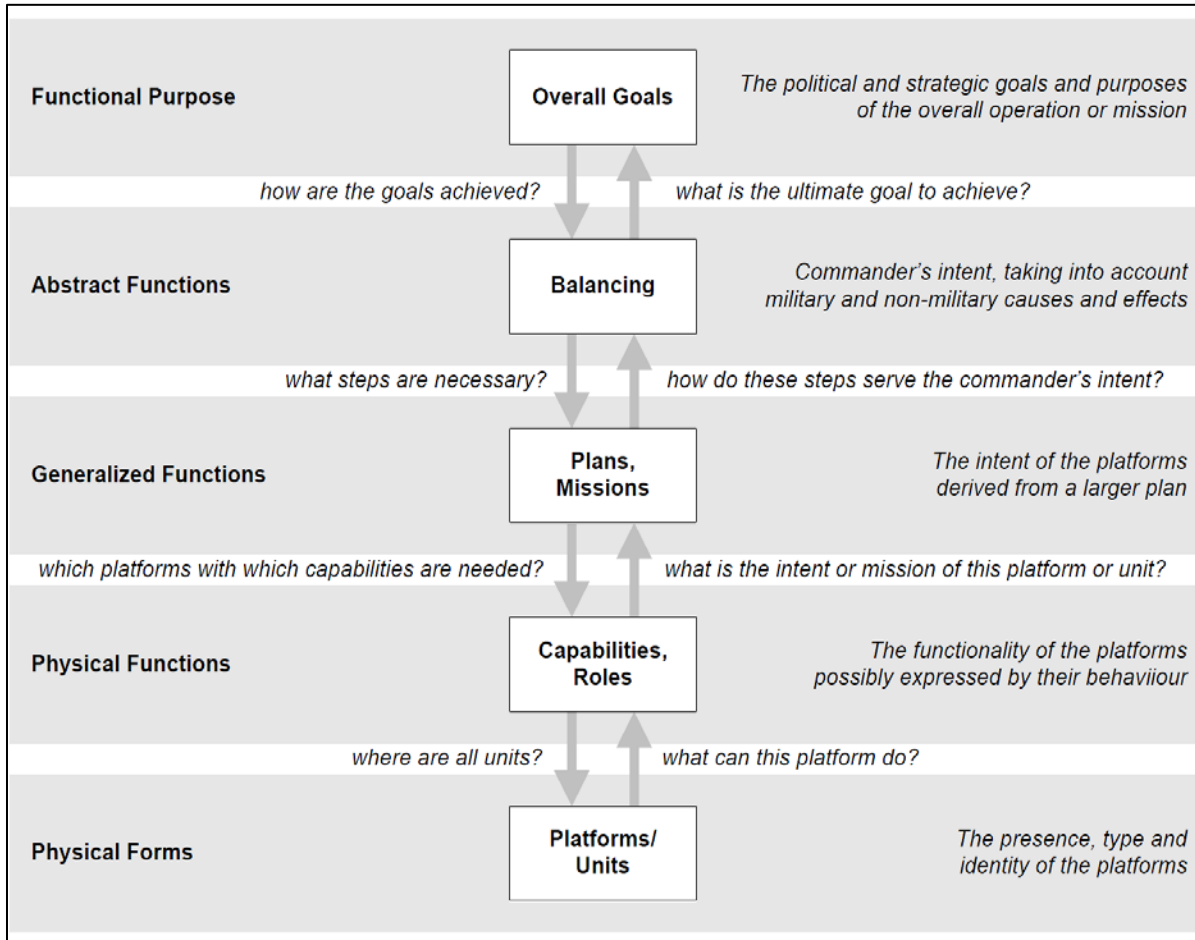


Figure 1. Abstraction Hierarchy of military command and control information.

Lower levels of abstraction consist of physical information (e.g., physical form level and physical functions level), higher levels of abstraction include abstract information (e.g., Functional purpose level and abstract functions level). For example, lower abstraction levels in a naval command and control system include sensor information like heading and speed indicators, higher abstraction levels include behavioral intention and possible actions of entities.

Structuring information as an abstraction hierarchy can typically be done along two dimensions: one is the abstraction hierarchy in information representation, the other dimension is the system's whole/part decomposition in which system parts aggregate from single components (physical) to the whole system (functional). This two dimensional space is referred to as the *abstraction decomposition space*.

Concerning an effective representation of a system's status data, a hierarchical representation of system parts can be represented in layers; consistent with the abstraction hierarchy. For example, Treurniet, van Delft, and Paradis (1999) described how different processes in a naval scenario can be modeled in a *tactical abstraction decomposition space* (TADS). Basic 'lower order' information like coordinates and speed measures of surrounding ships is supplemented with abstract information like law breaking indication or lane adherence. Advanced technological systems are able to automate data processing steps, taking over a reasoning step in information processing which is currently done mostly by human operators.

In this study, we use a TADS based on Arciszewski and De Greef (2011). In this concept, the system automates information processing steps in order to generate higher order (abstract) information. In that way, information is provided at different levels, matching three levels in the abstraction hierarchy (physical form level, physical function level, and generalized function level). We compare this TADS implemented interface with an interface that represents only physical form level information. Important questions in this study are what are the effects of TADS on the operator's comprehension, and more importantly what is the effect on the communication within a team of operators?

Situation Awareness

As is stated in the introduction, operators use a technological system in order to achieve a situation awareness (SA). Endsley (1995) defined situation awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". Endsley noted that SA should be distinguished from situational assessment, which is the process in order to achieve SA. Situation assessment is defined by a three level concept. Level 1 is defined by *perception*, the first step in achieving SA. Level 2, as the next step in situation

assessment, is characterized by *comprehension*, and level 3 identifies the *projection* of the future status. Level 3 is considered necessary for full achievement of SA, but not sufficient. That is, a mental model consisting of level 3 (projection of future status) still could lack perception (level 1) and comprehension (level 2).

In our study, we examine the SA in order to assess the operator's understanding of a tactical situation at different levels of abstraction. The main question here is, does SA increase when abstraction hierarchy is applied in the representation of information in a domain specific system (TADS)?

The issue of how SA should be measured has been discussed for years in literature (a summary of different concepts of SA measurements is attached in appendix A). An objective and commonly used SA measurement method in infantry, aviation, navy, and air traffic control is the situation awareness global assessment technique (SAGAT) by Endsley (1995). The SAGAT includes statements about the situation or the system's status. The participant assigns each statement as true or false. Based on the SAGAT, McGuinness (2004) developed the quantitative analysis of situational awareness (QUASA) technique. This technique differs from SAGAT in that the QUASA uses four alternative forced choice questions instead of true/false statement questions. Also a confidence bias is part of the QUASA technique. The confidence bias is an indication of the perceived SA and differs from actual SA in which the operator is either correct or incorrect about a statement. SA can be measured 'real-time', during an interruption of the task, or during a post-test following the task. In this study, the experiment is based on a simulated scenario in which it is possible to interrupt trials for SA questions, therefore we choose to use the interruption method to assess the SA on multiple moments in time. For the current study, we used the QUASA technique without correcting for confidence bias since we are interested in the actual SA as a mental model, independent of cognitive processes like decision making.

Communication

Operators rarely work alone, on the contrary, a team of operators usually work together with a common system. Collaboration is an essential part in performing this task. Since a mental representation is assumed to be related to the abstraction hierarchy in the system's information representation, it seems plausible that verbal communication within the team is influenced as well.

Much research focused on team communication, which is defined as information transfer between team members in collaboration settings. When task-relevant content in communication is measured as a fraction of the total amount of information transferred (the information density), Gorman et al. (2003) found that team performance increases with a larger fraction of task relevant content. Results found by Orasanu (1990) and Foushee and Manos (1981), show that the number of communication acts in aircrews is related to task performance. Better performing teams tend to communicate more frequently as compared to teams performing lower on the task. However, Thornton (1992) demonstrated that more communication did not induce higher performance, on the contrary, the number of errors appeared to be positively correlated with the number of statements made in communication. McKendrick et al. (2011) suggest that an increase in workload and cognitive demands might hinder the operator's ability to effectively communicate to collaborators, while increased communication between team members may be another source of cognitive demand (Bowers, Braun, & Morgan, 1997). A plausible explanation for these inconsistent results is demonstrated by Mosier and Chidester (1991) who found a negative relation between performance and the amount of communicated information. The authors reasoned that SA (as a predictor for performance) only affects the communication when highly demanding situations are apparent.

In addition, it is noteworthy to mention the findings of Bowers et al. (1998) that consistency in communication sequence (e.g., commanding statements are consistently replied by confirmation statements etc.) was found higher in better performing teams as opposed to worse performing teams. The authors suggest that “losses attributable to mismatches between expectations and actual communications may be reduced through the standardization of communication sequences.” In line with this reasoning, we foresee this plausible artifact by analyzing standardized communication, namely the standard protocol for verbal situational reports in military operations.

Communication can be analyzed with different methods, including the flow of communication (i.e., who is talking to who, when, how often one is communicating) and content analysis (i.e., what and how much information is transferred, how is the information related, what and which themes/categories can be distinguished). The assessment of communication in our research is context based analysis, namely *what* information is transferred and not *how* the communication occurs. Specifically, we are interested in what operators report to perceive. To analyze content of communication, we used keyword spotting and statement coding to determine the presence of certain themes or categories. Namely, the presence of higher order (abstract) information, lower order information and the presence of attribution statements about classification, identification, behavior and mission.

The focus in this study is to examine the presence of abstract information in team communication when abstraction hierarchy is applied in a domain specific system (TADS). That is, do operators communicate in higher levels of abstraction when abstraction hierarchy is applied in information representation? Is the situation communicated at a higher density when abstract information is represented in the system’s interface?

Along with these measures of communication and comprehension, we measure workload in order to assess the subjective mental workload. We expect that participants

experience the workload the same or lower with TADS implemented interfaces as the task description is the same in both conditions.

Summary

By applying concepts of EID, environmental information can be presented in a system more effectively. By using an abstraction hierarchy in the information representation, a better fit to the operator's mental representation can be realized. We are interested in the effect of abstraction hierarchy on communication. That is, how is communication within a team of operators affected by the information representation of the system?

Results from past research showed that SA increases with abstraction hierarchy, but the effect of mapping information at different levels to system parts (TADS) on SA and communication is not examined yet. It is known that communication is related to abstraction hierarchy in information representation and SA, but content based studies are not conducted yet to examine the relation between communication and abstraction decomposition spaces. Questions in this research are: do operators communicate in higher levels of abstraction when abstraction hierarchy is applied in information representation? Is the situation communicated at a higher density when abstract information is present in the system's interface? Does SA increase when abstraction hierarchy is applied in the representation of information?

To answer these questions, an experiment is set up. In this experiment, abstraction hierarchy is applied in an interface of the Dutch naval combat management system. The implementation of the abstraction hierarchy is drawn and elaborated to the naval specific domain by Arciszewski and De Greef (2011). In the next chapter, the TADS implementation is detailed by describing the system's interface. As the manipulation variable, we compare the TADS implemented interface with a conventional interface in which no abstract information is represented.

SA measures are scaled as the number of correctly assigned entities in a tactical situation. Communication is analyzed for its content to address the abstraction levels and information density is a measure for efficiency. By distinguishing higher and lower order information, we examine how communication at higher and lower order is affected by the implementation of TADS in information representation.

Interface and Team Communication in Naval Domain

The overall goal of a socio-technical system is to create awareness about the tactical situation in order to prepare for action to and anticipation on the situation. In a typical command and control room, operators are assessing the situation by making conclusions about behavior and missions based on information that is available. The situation is defined as the total picture of physical environment (e.g. the landscape) plus the knowledge about the objects in it (e.g. characteristics of known vehicles and vessels). These objects, referred to as *tracks*, comprise kinematic data like speed, heading, and altitude. Tracks also can be classified as airplane, ships, commercial air, etc. Finally, tracks can be identified as friendly, assumed friendly, neutral, suspect, and hostile. The operator analyzes this data in order to make conclusions about the behavior and possible missions of the track. The behavior and mission of a track, and possibly the intentions of the organization behind it is thereby assessed in an efficient and structured manner.

TADS Design

For this naval specific setting, the TADS is applied to the system that is known in military domain as the combat management system (CMS). The proposed interface is focused on helping the operator building the situational picture. That is, providing the operator with indications of environmental information (i.e., situational knowledge) at abstract levels. That is, higher order information like behavior and possible missions of surrounding entities. The CMS collects and displays information that is derived from sensors (e.g. sonar and radar), identification systems (e.g. IFF, AIS), and manually inserted information (e.g. surface maps, airways, and sea lanes). In the next sections, the implementation of TADS is explained by describing the systems components. In the system, possible behavior and missions are estimated and presented in the system's interface (referred to as the system's advice). The algorithms and criteria for the system's estimates can be found in Arciszewski and De Greef

(2011). Information apparent in the CMS in the conventional and TADS implemented displays are shown in Table 1. The most used attributes for these processes are identification response, sensor observations, and adherence to geographical areas like airways, sea lanes, and fishing grounds. The classification and identification process take place in the lowest level of the abstraction hierarchy, the physical form level.

Table 1. *Information apparent in the CMS for the conventional and TADS implemented displays.*

Indicator	Apparent in CMS	
	Conventional	TADS implemented
Tactical display		
Geographical information	x	x
Kinematic information	x	x
Track detail display		
Classification	x	x
Identification	x	x
Kinematic information	x	x
Behavior		x
Mission		x
Reasoning sections		x
Behavior/mission display		
Mission		x
Bearing and CPA		x

In this research, we applied the concept of TADS into a Dutch CMS. The conventional interface is complemented with the system's estimations about behavior and possible missions, a reasoning section. The interface of the conventional CMS includes two displays: a tactical display representing a geographic overview and a track detail display representing information about a selected track. The TADS interface is complemented with a behavior/mission display overviewing the bearing and closest point of approach (CPA) for the neighboring tracks. In the tactical display (Figure 2), a geographic overview of the tracks on the map are shown. This display provides low level information, including positions, track numbers and kinematic information.

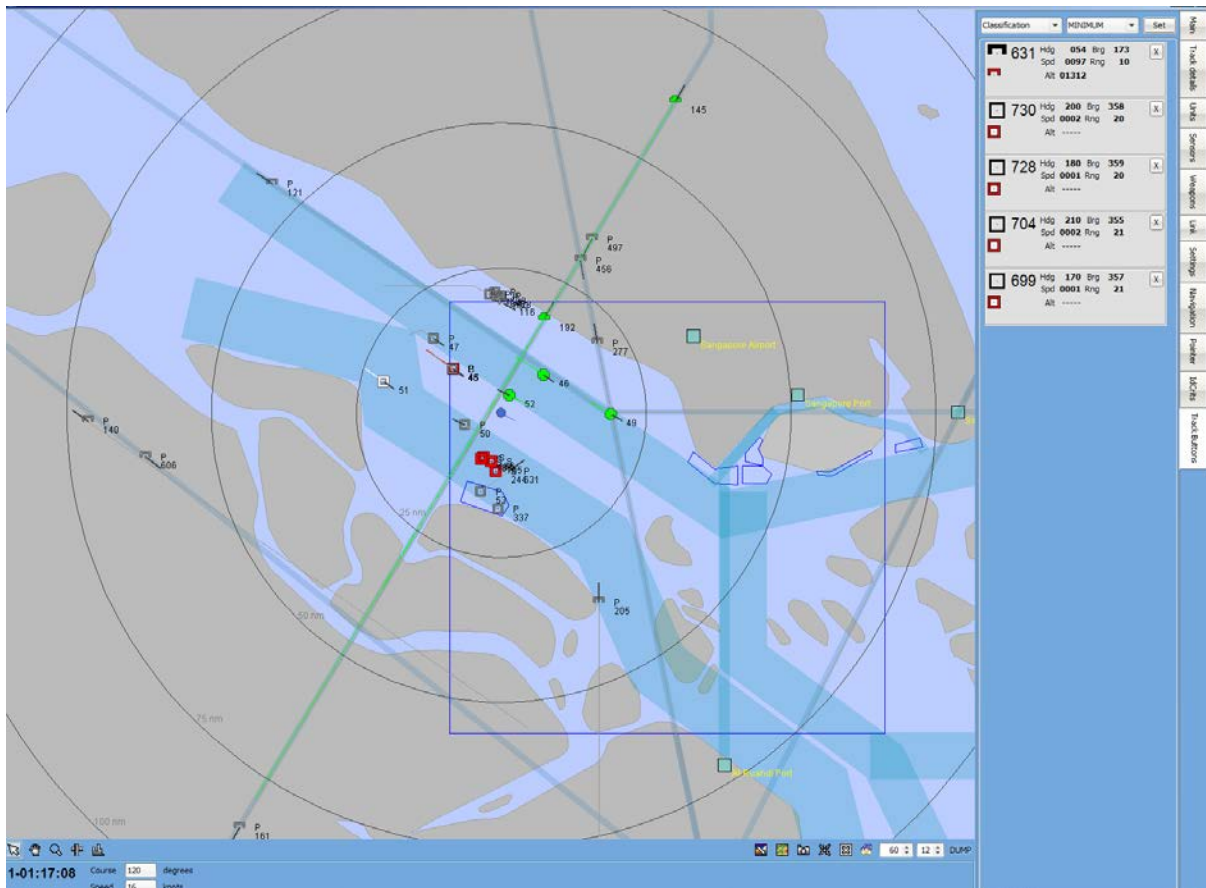


Figure 2. Tactical display.

In the track detail display (Figure 3), information about the selected track is presented. Indicators of physical function level are located in section A. Classification indicates the general and specific type of vessel or vehicle (e.g. sea, containership) and its entity (e.g. a frigate). Also the assigned identification is presented. This indicator consists of five options (hostile, suspect, neutral, assumed friendly, friendly) and accompanied by the system's advice (highlighted yellow). In section B, possible behaviors and missions are listed. These indicators are at abstract levels, namely the generalized function level. The system's advice are the possible behavior and missions indicators highlighted yellow. For example, 'recreational flight' as behavior and 'smuggling' as a possible mission. Section C and D represent the reasoning section for classification and identification. That is, the information that was evident to conclude a certain possible class, identity, behavior or mission can be assigned to the track. In section E, basic sensor information is presented. This is the system

part level in the abstraction hierarchy and provides elementary data including heading, speed, altitude, CPA, bearing, and history graphs of height, speed, and CPA.

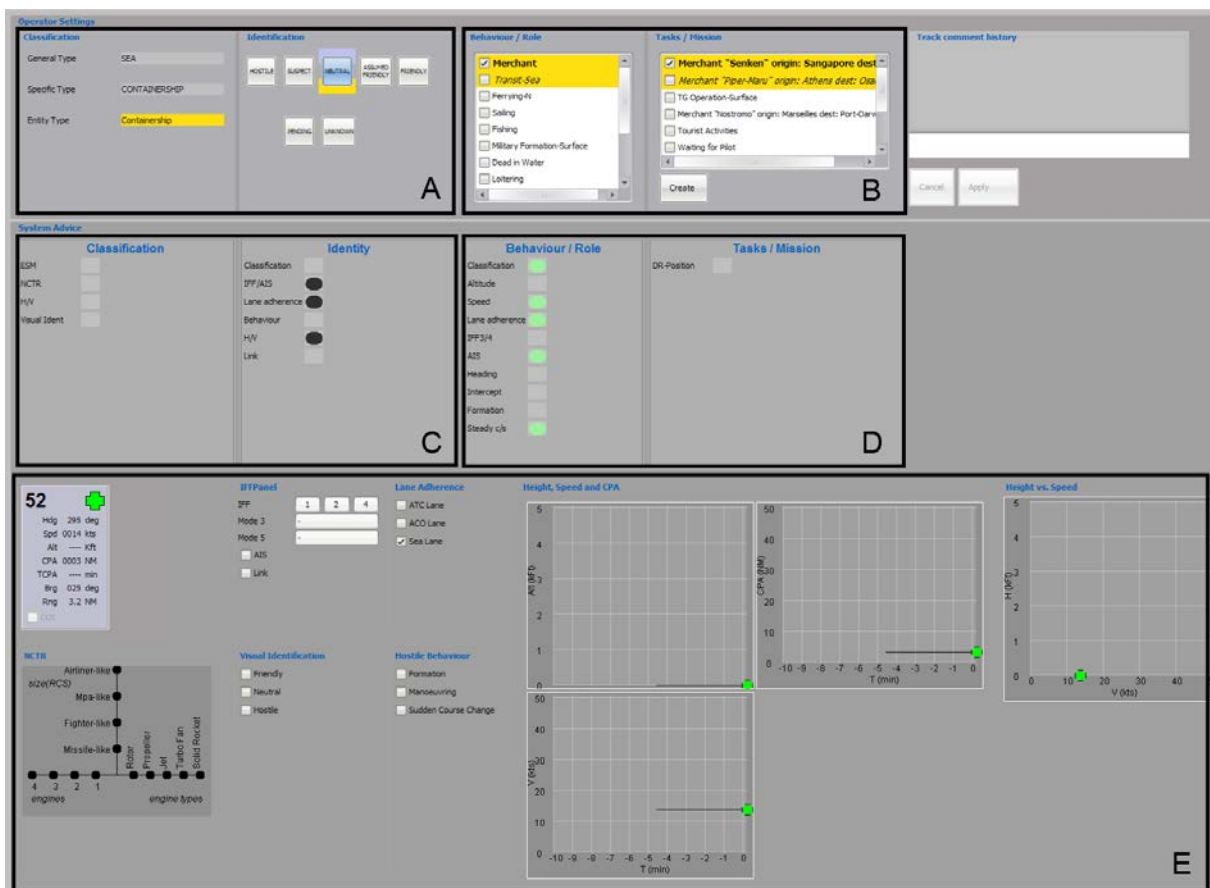


Figure 3. Track detail display, letters indicate sections of indicators at lower order level (A and E), higher order information added in a TADS (B, C, and D).

The behavior/mission display represents a behavior and mission overview (Figure 4). In this two dimensional space, with bearing on the horizontal axis and time on the vertical axis, tracks are positioned (section B). These tracks are color coded (identity) and are connected to a dashed line, representing its direction (approaching or receding). Along the left side (section A) and right side (section C) of this overview, assigned and suggested sea missions (left) and air missions (right) are presented. This third display consists of only functional information since only abstract information (mission and behavior) is presented. This display essentially represents all the information required for a complete situation assessment.

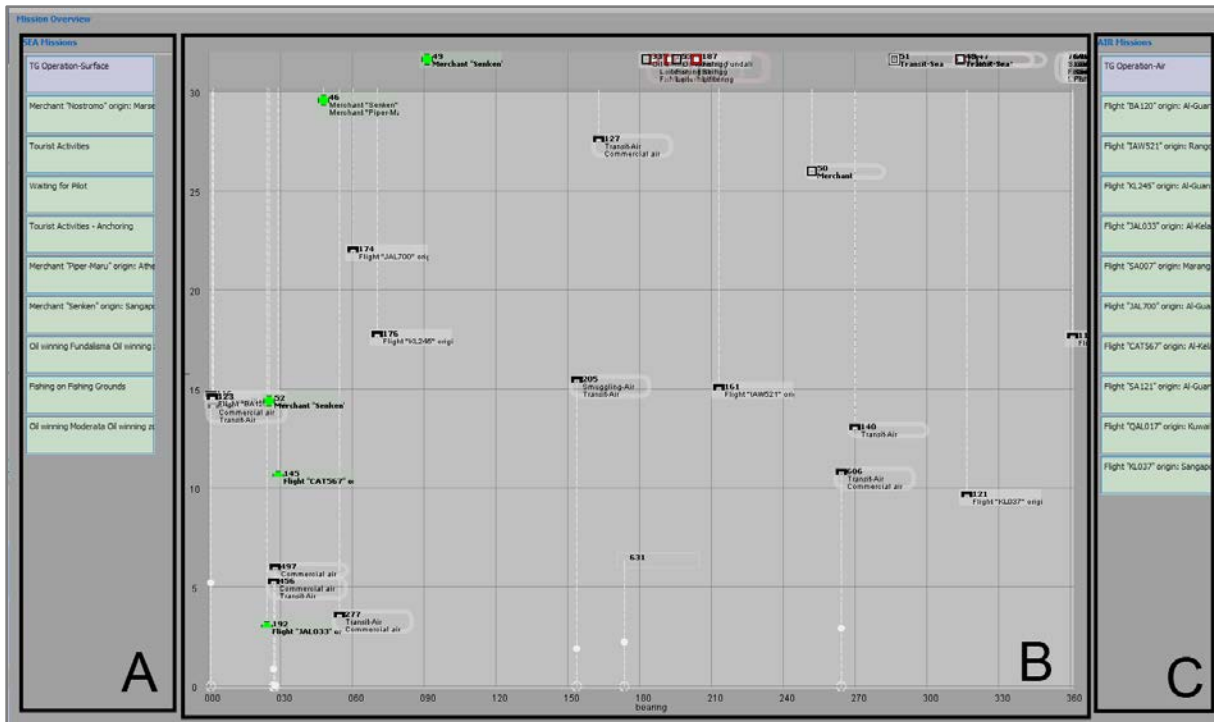


Figure 4. Behavior/mission display representing tracks plotted over bearing and CPA (section B) accompanied by possible sea (section A) and air (section C) missions.

Team Communication

The task is described for one operator, but other collaborators (on either the same or another location) perform a similar task. Therefore, resulting observations (situation assessments) have to be communicated effectively in order to plan further actions. In a typical command and control, the situation is communicated verbally and is referred to as a situational report (sitrep for short). When a sitrep is performed, three subjects are described (if possible): what is observed, what does that entail and what is the proposed plan of action.

Note that the focus of this research is to address the SA and communication in a tactical operation. Decision making and planning and action in the operation is beyond the scope of this paper.

Method

In the experiment, naval participants were acting as operators in a simulated command and control environment. Six comparable tactical situations have been simulated, these situations are referred to as scenarios. When participants were carrying out the scenario run (i.e., a trial), three moments of interruption have been introduced for sitreps and SA questionnaires. Information was represented in either a conventional CMS (baseline condition) or a CMS with TADS applied (TADS condition). Mean scores of SA and communication were compared in order to address the effect of the implementation of TADS. This chapter describes the task, experimental design, measures chosen for SA and communication, procedure and approach of analysis.

Participants

Sampling was done by random assignment of officers from the Royal Netherlands Navy (RNLN). A sample of 12 participants were taken. 11 males and 1 female were aged between 34-57 years. All participants had extensive experience using naval CMSs. At the time of the experiment, participants were either involved in CMS development or in training of officers at the institute of operational education (OPSCHOOL). Each trial was carried out by two participants: one participant was acting as the first operator, the other participants was acting as the second operator.

Task

The task of the operator is to build a situational picture by assigning identity, behavior, classification, and missions to tracks in the system. Participants were briefed to carry out an UN mission concerning prevention of smuggling activity (detailed in appendix B). To achieve this task, the participants were required to use the information from the CMS and were able to ask for clarification from external sources (e.g. making radio contact to the coastguard requesting identification of a certain track). It was not allowed to make notes except for

assignment of class, identity, behavior, and mission into the CMS. On three specified moments in every trial, the participant acting as the first operator was required to perform a sitrep. The protocol for a sitrep was to report three steps: what is observed, what does it mean, and what are future recommendations for action. The task of the second operator was keeping track of the situation by observing the systems displays on the same CMS. The participant was not allowed to make adjustments in the system. This task was performed half of the time. The other half of the time the second operator was turned away from the displays and was required to perform different tasks (puzzles were provided). This alternation of the tasks is implanted in the procedure to create the realistic situation where collaborators are frequently multitasking and not always aware of the tactical situation.

Manipulation

Two conditions were compared, specifically the condition where the displays of the CMS is supported by TADS (i.e., added indicators of behavior and mission, the reasoning section, and the presence of the behavior/mission display), and the baseline condition where TADS support was not present (i.e., the tactical display and the track detail display without system view and reasoning section). Six comparable scenarios were created and were based on the same situation as described in Appendix B. Couples of two participants carried out four of the six scenarios, two trials in baseline condition and two trials in TADS condition. The order of which scenarios appear and which scenario runs in what condition was counter balanced. Participants were randomly assigned to a role (first or second operator) and reversed roles according to a balanced randomized schedule.

Measures

Multiple measurements were used for SA, workload, and communication constructs (an overview is shown in Table 2). SA and workload were measured by questionnaire, whereas communication pattern is measured by nine scales and derived from textual

transcripts of the sitreps. Two communication pattern measures were based on word counts, namely total word count and density. Two measures were based on keyword spotting: abstraction level 1 presence (lower order information) and abstraction level 2/3 presence (higher order information). Five measures were based on hand coded statement counts divided into six categories: kinematic, class, identity, behavior, mission, unrelated statements.

Table 2. Measures with accompanying method and construct used in the experiment.

Measure	Method	Construct
Situation Awareness (SA)	Probe questionnaire on intermitted moments in operation	Four Multiple choice questions querying activity, identity, behavior, and mission accompanied with a five level confidence scale
Workload	Questionnaire form after each trial	Uni-directional scale, ranging from 0 to 150.
Total word count	Automatic word count	Total number of words in a sitrep
Density	Automatic word count	The proportion of unique words in relation to the total number of words
Keyword presence abstraction level 1	Automatic keyword spotting	The proportion of Level 1 in the sitrep in relation to the total number of words spoken
Keyword presence abstraction level 2/3	Automatic keyword spotting	The proportion of Level 2/3 in the sitrep in relation to the total number of words spoken
Statement presence	Statement coding	The proportion of a category (5) specific statement in relation to the total number of statements

SA measures were based on the quantitative analysis of situational awareness (QUASA) method of McGuinness (2004). 36 questions concern the state of 12 specific tracks. These tracks were presented along with the other tracks in the recent screenshot of the tactical display. Four multiple-choice questions were asked about the Identity, Behavior, Activity, and Mission (listed in appendix C). Answers can be correct or incorrect compared to the current situation. A SA question is for example: “What is the Identity of track 315?”. The participant could use the provided tactical display screenshot to lookup the specified track. The SA measure was calculated by the total number of correct answers of all 36 questions in each

trial. We hypothesize that the mean SA score is higher in the TADS condition compared to the baseline condition.

Workload is measured directly at the end of each trial by the *rating scale mental effort*, RSME for short (Zijlstra & Van Doorn, 1985). The RSME form was a scale ranging from 0 to 150. We expect that mean workload in the TADS condition is equal or higher than the mean workload in the baseline condition.

The total word count is measured in absolute numbers by counting the number of spoken words. Counts were performed by computer and based on transcripts. Total word count is expected to be equal in both TADS and baseline conditions.

The density is measured by counting the unique words used in the sitrep and divided by the total word count:

$$Density = \frac{Number\ of\ unique\ words}{Total\ word\ count}.$$

The density is expected to be higher in the TADS condition compared to the baseline condition. That is, a relative higher diversity in words is expected to be found.

With keyword spotting, words were specified to a specific category. Two measures for low and high abstraction level presence were defined: abstraction level 1 and abstraction level 2/3 (keywords are listed in Appendix D). The abstraction level 1 and level 2/3 presence (for $X=1$, $X=2/3$) were calculated by dividing the word count of the level specific keywords by the total word count:

$$Keyword\ level\ X\ presence = \frac{Keyword\ level\ X\ count}{Total\ word\ count}.$$

The presence of level 1 keywords is expected to remain equal over the TADS and baseline conditions since level 1 information is present in both baseline and TADS condition. Level 2/3 presence is expected to increase with the TADS condition.

All statements in sitreps (i.e. expressions corresponding to notifications and informational messages, sentences) have been coded into 6 categories: kinematic, class, identity, behavior, mission and unrelated statements. Coding is done manually by human experts and is detailed in the analysis section of this chapter. Presence of kinematic, class, identity, behavior, and mission statements have been calculated by the count of category specific statements divided by the total count of statements found in the sitrep:

$$\textit{Statement X presence} = \frac{\textit{Number of statement X}}{\textit{Total statement count}}.$$

We hypothesize that a higher presence of behavior and mission statements will be apparent in the TADS condition compared to the baseline condition while kinematic, class and identity statements are expected not to differ.

Apparatus

A command and control room is realized by setting up a workstation for each participant (for the first operator and the second operator role). The workstations consisted of a panel of the tactical display, track detail display, and the behavior/mission display. Controls consisted of a keyboard and a mouse (a workstation is shown in Figure 5). In the same room, a second workstation was located on which the SA (Figure 6) and RSME (Figure 7) questionnaires were filled out.

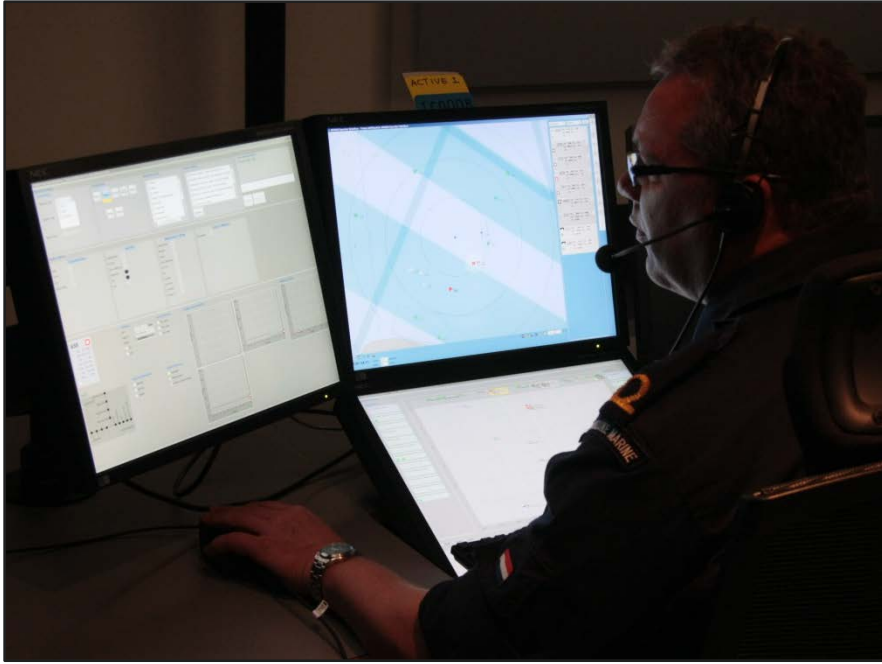


Figure 5. Participant behind a work station.

Participants were positioned in physically, auditory, and visually separated rooms. The same scenario ran on both operator stations, a third station is located in a separate room from which the scenario is monitored by the experimenter. From this room, the experimenter acts as the ‘external world’, i.e., representing collaborators and other contacts that can be contacted by radio. Radio contact between the first and the second operator was enabled when the sitreps were given. Sitreps have been recorded and transcribed into textual formats by computer using Nuance Dragon Naturally Speaking (2011) and corrected manually afterwards. The resulting transcripts of the sitreps formed the dataset for analysis.

What is the identity of

identity:

Track 109:

Track 244:

Track 239:

confidence:

very unsure very sure

Figure 6. Screenshot of SA questionnaire.

Inspanningsschaal

Wilt U door middel van het zetten van een streepje op onderstaande lijn aangeven hoeveel inspanning het U heeft gekost om deze taak uit te voeren

150

140

130

120

110 ————— ontzettend inspannend

100 ————— heel erg inspannend

90 ————— erg inspannend

80

70 ————— behoorlijk inspannend

60 ————— tamelijk inspannend

50

40 ————— enigszins inspannend

30 ————— een beetje inspannend

20

10 ————— nauwelijks inspannend

0 ————— helemaal niet inspannend

Figure 7. RSME questionnaire, participants indicate their subjective workload by positioning a slider on the vertical scale.

Procedure

In each trial, a scenario was run for 30 minutes. The first operator was executing his task while the second operator alternated between the observing task and the puzzle task every five minutes. A diagram for a whole trial is given in Figure 8. Every 10 minutes, the scenario was paused and the screens were blanked. During the intermission, first the SA questionnaires were filled out, followed by the sitrep performed by the first operator to the second operator. During the sitrep, radio contact and the displays were enabled again. The scenario was still paused as long as the sitrep was performed. A maximum duration of two minutes was set for sitreps.

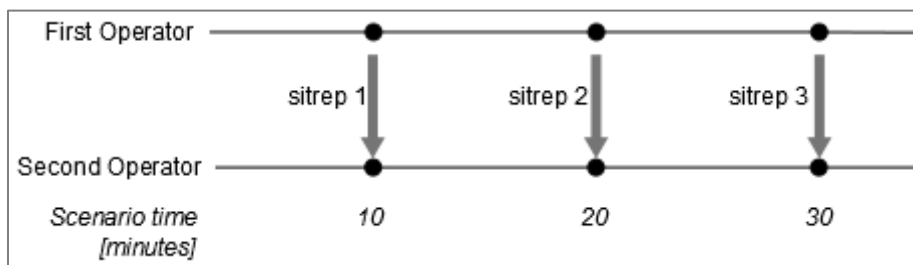


Figure 8. Timeline and procedure in a trial.

Experiments started with a briefing in which the situation in the scenario was explained instructions were given. After the briefing two training trials were executed in order to get the participants used to the situation (i.e., the new interface and the tactical situation and mission of the scenario). At the end of the trials, couples attend to a general discussion to give comments and feedback about their subjective experience with the interface.

Analysis

For keyword spotting, a set of abstraction level specific words have been defined. As a starting point, keywords were collected from sitreps in a previous naval experiment (Essens et al., 2011). Firstly, commonly used words were listed. Secondly, five human raters assigned each word to a specific category (i.e. either abstraction level 1 or 2/3). Then, selection of keywords for each abstraction level category was done by filtering the words of which more than 80% of the raters agreed with. Words that had an inter rater agreement of 60% or less

were rejected. Finally, the resulting set contained 101 keywords: 36 level 1 keywords and 65 level 2/3 keywords.

To determine the inter rater reliability of the statement category scales, three expert raters have coded three sample sitreps. A generalized kappa (Fleiss, 1981) indicated that overall inter rater reliability was moderate for all statement categories (kappa = .48). The kappa was considered high for the categories kinematic (.57), class (.61), and identity (.53) statements. A moderate kappa have been found for the category behavior statements (.49) and a low inter-rater reliability for the mission statement category (.17).

Analysis is based on data from the first operator role, the participant who was giving the sitreps. All analyses were done using within participants repeated measures (analysis of variance, ANOVA). Comparison of Baseline versus TADS conditions is based on the mean scores of the measures of the three sitreps in each trial. Every condition of each measure therefore consists of 12 data points.

Results

Means and standard deviations for both Baseline and TADS condition are presented in Table 3. Mean SA scores in the TADS condition and the baseline condition have been compared and tested. Results are plotted for each condition for each measure in Figure 9.

Table 3. Mean (*M*) and Standard Deviations (*SD*) and difference test values (*F*) and significance levels (*p*) of SA, Workload, and Communication Measures for the baseline and TADS condition (*n*=12).

	Baseline		TADS		<i>F</i>	<i>p</i>
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)		
SA	22.17	(4.93)	24.17	(4.13)	3.67	.082
Workload	68.83	(15.4)	66.33	(12.6)	0.64	.440
Total word count	173.1	(56.0)	177.1	(52.6)	0.20	.660
Density	.612	(.076)	.619	(.066)	0.22	.647
Proportion level 1 keywords	.069	(.023)	.061	(.016)	2.55	.138
Proportion level 2/3 keywords	.040	(.019)	.046	(.014)	3.46	.090
Kinematic statements	.179	(.055)	.171	(.046)	0.21	.655
Class statements	.107	(.058)	.105	(.035)	0.02	.905
Identity statements	.115	(.050)	.124	(.064)	0.19	.669
Behavior statements	.200	(.068)	.204	(.051)	0.13	.721
Mission statements	.088	(.046)	.097	(.030)	0.46	.510

Results from the repeated measures analysis of variance showed that the mean SA score in the TADS condition did not differ from the baseline condition significantly ($F_{(1,11)} = 3.67, p = .082$) but indicated a trend toward the hypothesis that the SA is higher in the TADS condition: $M_{\text{baseline}} = 22.17, M_{\text{TADS}} = 24.17$. The mean workload in the TADS condition did not differ significantly from the mean workload in the baseline condition ($F_{(1,11)} = .64, p = .440, M_{\text{baseline}} = 68.83, M_{\text{TADS}} = 66.33$). When the mean total word count is compared for the TADS and Baseline condition, no significant difference was found ($F_{(1,11)} = .20, p = .660, M_{\text{baseline}} = 173.1, M_{\text{TADS}} = 177.1$). Also the mean scores of density in the TADS and baseline conditions appeared not to differ significantly: $F_{(1,11)} = .22, p = .647, M_{\text{baseline}} = .612, M_{\text{TADS}} = .619$. The mean proportion of level 1 keywords did not differ significantly: $F_{(1,11)} = 2.55, p = .138, M_{\text{baseline}} = .069, M_{\text{TADS}} = .061$. The mean proportion of level 2/3 keywords however showed a nearly significant difference: $F_{(1,11)} = 3.46, p = .090$. Participants tend to use more

level 2/3 keywords in the TADS condition as opposed to the baseline condition ($M_{\text{baseline}} = .040$, $M_{\text{TADS}} = .046$). For the mean proportion of statements, no significant difference was found between the baseline and TADS condition (kinematic statements: $F_{(1,11)} = .21$, $p = .655$, class statements: $F_{(1,11)} = .02$, $p = .905$, identity statements: $F_{(1,11)} = .19$, $p = .669$, behavior statements: $F_{(1,11)} = .13$, $p = .721$, and mission statements: $F_{(1,11)} = .46$, $p = .510$).

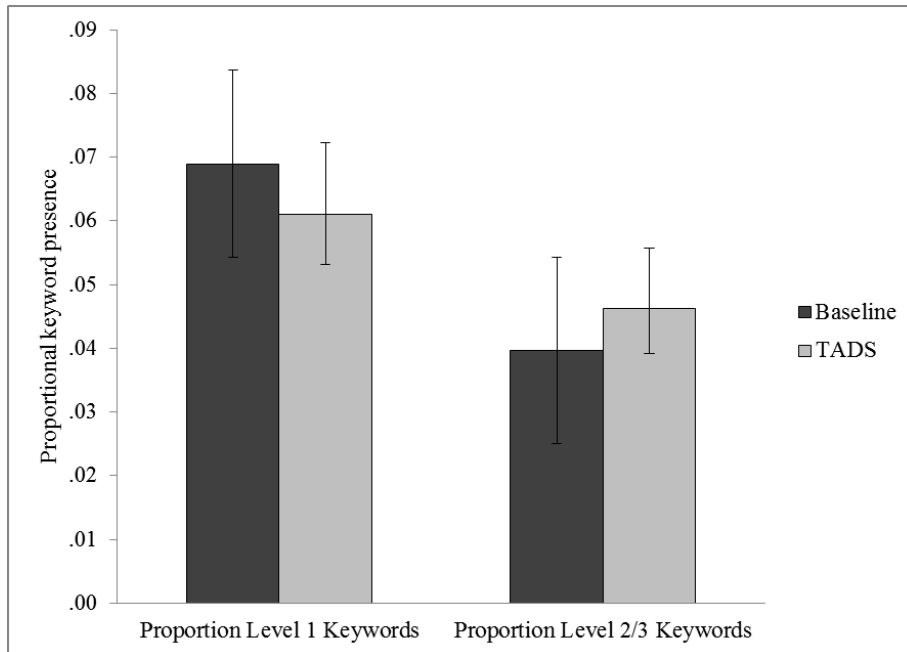


Figure 9. Mean keyword presence and standard deviations for level 1 and level 2/3 keywords.

Discussion

Operators of socio-technical systems have to deal an increase in amount and complexity of information. Risks of misperceiving and incorrectly adapting to change and novel information increase when operators are executing the task of assessing the system's presented situation. By applying concepts of EID (i.e., information representation as an abstraction hierarchy, mapped out in an abstraction decomposition space), environmental information can be presented in a system more effectively. By using an abstraction hierarchy in the information representation, a better fit to the operator's mental representation can be realized. Our general question is what is the effect of abstraction hierarchy in information representation in a socio-technical system on the operator's comprehension and communication in a team of operators when verbal situation reports are given?

We applied abstraction hierarchy in a naval combat management system where system parts represent information on three levels of abstraction hierarchy. To measure the communication, we looked at the proportional presence of lower order information (keyword level 1 presence) and higher order information (keyword level 2/3 presence) apparent in communication. Also, we looked at relative variation of terms used in communication. Additionally, we looked at the amount of correctly perceived task specific elements presented by in the system in order to assess the situation awareness.

We expected – with an abstraction hierarchy in information representation – that in the communication, higher order information is more present in communication while lower information remained the same, and that there is a higher variation in spoken words. Also, a higher situation awareness was expected while workload remains at least the same, or even might lower with abstraction hierarchy in the information representation.

We found that the amount of lower order information in communication remained the same, but higher order information did not increase significantly with the application of abstraction hierarchy in information representation.

Also the relative variation of words used in communication (the density) did not increase significantly. The finding by Gorman et al. (2003) that team performance is related to communication density could not be complemented by significant effects of abstraction hierarchy.

Against our expectation, the situation awareness did not increase when abstraction hierarchy is applied. However, what was expected, is that the workload remained the same. The suggestion that SA and communication are negatively related only in highly demanding situations (Mosier and Chidester, 1991), is not confirmed by this study since the mean workload was around the middle of the scale (between rather effortful and pretty effortful), assuming that there was highly demanding situation apparent. These results are representative for the application in naval settings, we cannot draw general conclusions for other domains with certainty since information in different domains could be of different nature, making the application of abstraction hierarchy to a specific domain subject to the nature of the information.

Information was available on three levels: physical form, physical function, and generalized function level. These levels might be semantically too close to differentiate between them. The interface might not be convincing enough, that is, abstract indicators have not been evident (enough) to the participant. It is not just simply take over the systems suggestions, but also the arguments of why tracks tend to show a certain behavior or act according to a certain mission. To address the track's 'whole story', lower levels of knowledge still have to be obtained. Using this lower order information to conclude possible behavior and missions (higher order information) is the conventional way the participants

used to perform the task. Participants still could have chosen to work in the conventional way, ignoring the TADS additions. In the general discussion after the trials, some of the participants confirmed this pitfall. Measures of statements in the communication were based on categories from the interface (statements concerning kinematic information, classification, identification, behavior, and mission). As the inter rater reliability test pointed out, the categories of higher order information like behavior and mission were respectively moderate and low in inter rater reliability. It appears that the categories for higher order information are difficult to define by human experts, probably causing unstable and unreliable results.

Unlike other research in team communication, we focused on sitreps. Another form of communication, namely inter-operator communication could be related to the knowledge representation in the systems interface. Kanki, Greaud, & Irwin (1991) for example found that the sequence of statements is related to the performance and situation awareness of team members. That is, the consistency in statements and response statements (e.g., commands are consistently replied with a confirmation) is related to the performance and situation awareness of team members. In the setup we used in our experiment inter-operator communication in this way was not possible. Future research could address this topic in a setting where open communication is possible.

Although no significant results were found, a trend was found for higher order words and SA. Near significant results ($p_{SA} = .090$, $p_{level\ 1\ presence} = .082$) indicated a trend towards our hypothesis that higher order information in communication and SA is higher when TADS is applied. We suggest that similar experiments with a more explicit representation of abstraction hierarchy might obtain significant results towards our hypotheses. Also, this experiment is executed by experienced operators. An experiment with novice participants might result in different conclusions as novice participants are not biased by training instructions in their education.

References

- Arciszewski, H. F., & De Greef, T. E. (2011). A Smarter Common Operational Picture: The Application of Abstraction Hierarchies to Naval Command and Control. *16th ICCRTS: Collective C2 in Multinational Civil-Military Operations*, 31(0).
- Bennett, K. B., Posey, S. M., & Shattuck, L. G. (2008). Ecological Interface Design for Military Command and Control. *Journal of Cognitive Engineering and Decision Making*, 2(4), 349–385.
- Bowers, C. A., Braun, C., & Morgan, B. B. (1997). Team workload: its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, research, and applications* (pp. 85–108). Hillsdale, NJ: Lawrence Erlbaum.
- Bowers, C. a., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing Communication Sequences for Team Training Needs Assessment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(4), 672–679.
- De Greef, T. (2010). Evaluation of Adaptive Automation using an Object-Oriented Task Model in a Realistic C2 Environment. *ePartners for dynamic task allocation and coordination* (pp. 72–94).
- Dragon Naturally Speaking. (2011). Burlington, Massachusetts: Nuance Communications. Retrieved from <http://netherlands.nuance.com/>
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1), 65–84. Retrieved from <http://www.ingentaconnect.com/content/hfes/hf/1995/00000037/00000001/art00005>
- Essens, P. J. M. D., Lindenberg, J., Post, W. M., Streefkerk, J. W., Houttuin, K., Janssen, J. B., Rakhorst, M. L. W., et al. (2011). *Genetwerkte Commandocentrales en Staven: Casus Commandobrug Patrouilleschip*. Soesterberg.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. (J. Wiley, Ed.) *Wiley series in probability and mathematical statistics* (Vol. 2nd, pp. 38–46). Wiley. Retrieved from <http://amstat.tandfonline.com/doi/abs/10.1080/00401706.1974.10489195>
- Foushee, H. C., & Manos, K. (1981). *Information transfer within the cockpit: Problems in intracockpit communications* (pp. 63–71). Moffett Field, CA.
- Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., & Cooke, N. J. (2003). Evaluation of Latent Semantic Analysis-based measures of team communications content. *Proceedings of the*.
- Grootjen, M., Neerincx, M. A., & Veltman, J. A. (2006). Cognitive task load in a naval ship control centre: from identification to prediction. *Ergonomics*, 49(12-13), 1238–1264. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17008255>

- Kanki, B. G., Greaud, V. A., & Irwin, C. M. (1991). Communication variations and aircrew performance. *International Journal of Aviation Psychology*, 1(2), 149–162.
- Koriat, A., & Goldsmith, M. (1996). *Monitoring and control processes in the strategic regulation of memory accuracy*. *Psychological Review* (Vol. 103, pp. 490–517). American Psychological Association. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8759045>
- McGuinness, B. (2004). Quantitative analysis of situational awareness (QUASA): Applying signal detection theory to true/false probes and self-ratings. *Command and Control Research and Technology Symposium*.
- Mckendrick, R., Shaw, T., Saqer, H., Visser, E. D., & Parasuraman, R. (2011). Team Performance and Communication within Networked Supervisory Control Human-Machine Systems, (2010).
- Mosier, K. L., & Chidester, T. R. (1991). Situation assessment and situation awareness in a team setting. *Designing for everyone: Proceedings of the 11th Congress of the International Ergonomics Association* (pp. 798–800). London: Taylor & Francis.
- Orasanu, J. (1990). *Shared mental models and crew performance*. Princeton, NJ.
- Paradis, S., Treurniet, W., & Roy, J. (1998). Environment Perception Process in Maritime Command and Control, (September), 14–17.
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems Man and Cybernetics*, 13(3), 257–266. Retrieved from <http://www.carlosrighi.com.br/177/Ergonomia/Skills>
- Rasmussen, J. (1985). The Role of Hierarchical Knowledge Representation in Decisionmaking and System Management. *IEE Transactions on Systems, Man and Cybernetics*, 15(2), 234–243. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/cbdv.200490137/abstract>
- Sulistiyawati, K., Wickens, C. D., & Chui, Y. P. (2011). Prediction in Situation Awareness: Confidence Bias and Underlying Cognitive Abilities. *The International Journal of Aviation Psychology*, 21(2), 153–174. doi:10.1080/10508414.2011.556492
- Thornton, R. C. (1992). *the effects of automation and task difficulty on crew coordination, workload, and performance*. Norfolk, VA: Old Dominion University.
- Treurniet, W., Van Delft, J. H., & Paradis, S. (1999). Tactical information abstraction framework in maritime command and control in Modelling and analysis of command and control. *Proceedings of the Symposium of the RTO Studies Analysis and Simulation Panel*. NATO report RTO-MP-38 AC/323 (SAS) TP/12.
- Vachon, F., Lafond, D., Vallieres, B. R., Rousseau, R., & Tremblay, S. Supporting situation awareness: A tradeoff between benefits and overhead. , 2011 IEEE International

MultiDisciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support CogSIMA 284–291 (2011). IEEE.

Vicente. (2002). Ecological Interface Design: Progress and Challenges. *Human Factors The Journal of the Human Factors and Ergonomics Society*, 44(1), 62–78.
doi:10.1518/0018720024494829

Vicente, K J, & Rasmussen, J. (1992). Ecological Interface Design : Theoretical Foundations. *IEEE Transactions on Systems, Man, And Cybernetics*, 22(4), 589–606.

Vicente, K. J., & Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating “direct perception” in complex work domains. *Ecological Psychology*, 2, 207–250.

Wade, J. (1996). *Navy Tactics, Doctrine, and Training Requirements for Littoral Warfare*. Naval Postgraduate School Monterey, California.

Zijlstra, F. R. H., & Van Doorn, L. (1985). *The construction of a subjective effort scale*. Delft.

Appendix A. SA summary derived from internal communication

Both McGuinness (McGuinness, 2004) and Sulistyawati (Sulistyawati, Wickens, & Chui, 2011) argue that the measurement of actual SA should be combined with measurement of self-perceived SA. They can be combined to a measure what is called calibration: “The principle of calibration concerns, in essence, the extent to which people are able to judge the correctness of their own observations or decisions. In other words, it assesses the degree of correspondence between self-perceptions of accuracy and actual accuracy as a proportion of correct responses (Koriat & Goldsmith, 1996). A well-calibrated judge is one who is highly confident about those responses that are in fact correct, and unconfident about those responses that are in fact incorrect. In a poorly calibrated judge, there is no systematic relationship between real and perceived accuracy. [...] In terms of situational awareness, a well-calibrated individual is one who has a high level of actual SA and correctly perceives this to be the case in his or her perceived SA. This can be assessed by correlating SA probe responses with confidence ratings in those responses. It is of course possible for an individual to be poorly calibrated with respect to SA. In the worst case, the individual is excessively overconfident - he has low actual SA but does not realize this and instead has high confidence in it. In this case, his decision-making is likely to be most error-prone.” (McGuinness, 2004). The difference between McGuinness and Sulistyawati is that the former uses true/false SA probes, whereas the latter uses a mix of different types of questions (open and true/false) similar to SAGAT. Both authors asked the participants for each probe question how confident they were. Sulistyawati used two confidence values (high and low), McGuinness used 5 confidence levels, however at the end of his paper he proposes to use two confidence levels, because it would yield better statistics. Another difference between the two authors is the analysis: McGuinness uses knowledge from the signal detection theory (Detection theory, receiver operating characteristic), he derives the sensitivity (the individual’s actual ability to

discriminate true signals from non-signals). He takes this to be the actual SA. He also derives the response bias, which "...specifies the setting of the participant's accept/reject criterion. A conservative bias reflects a tendency to reject whereas a liberal bias leans toward accepting." (Vachon, Lafond, Vallieres, Rousseau, & Tremblay, 2011). Sulistyawati on the other hand calculates the accuracy; the percentage of correctly answered probe questions. McGuinness questions this approach: "In probe-based techniques, SA has typically been assessed as the proportion of queries that are responded to correctly. While this seems an obvious statistic to use in terms of face validity, on its own it fails to provide a full picture of the subject's awareness because it confounds sensitivity and response bias. Does a low percentage correct reflect poor sensitivity or a highly conservative response strategy (responding selectively only when absolutely certain)? Discriminating between subjects' sensitivity on the one hand and response strategy on the other could be invaluable for understanding patterns in people's situation assessments." (McGuinness, 2004) Both authors calculate the calibration bias by subtracting the accuracy from the average confidence. In this case a positive calibration bias corresponds to overconfidence. McGuinness calls his probing method Quantitative Analysis of Situational Awareness (QUASA).

Appendix B. Scenario description

Six comparable scenarios were created around a United Nations resolution mission preventing smuggling of people, weapons, and narcotics. The goal of the operators is to compile a complete tactical maritime picture requiring identifying and addressing possible mission, role, class and identity for all contacts. A helicopter or unmanned vehicle is available but acts autonomously under control of the commando task group (CTG). The ship follows a course that was fixed per scenario. The scenarios were situated in the strait between Western-oriented Moderata and Islamic Fundalisma (see Figure 10). In Moderata, a number of Islamic groups are active enabling a Islamic revolution.

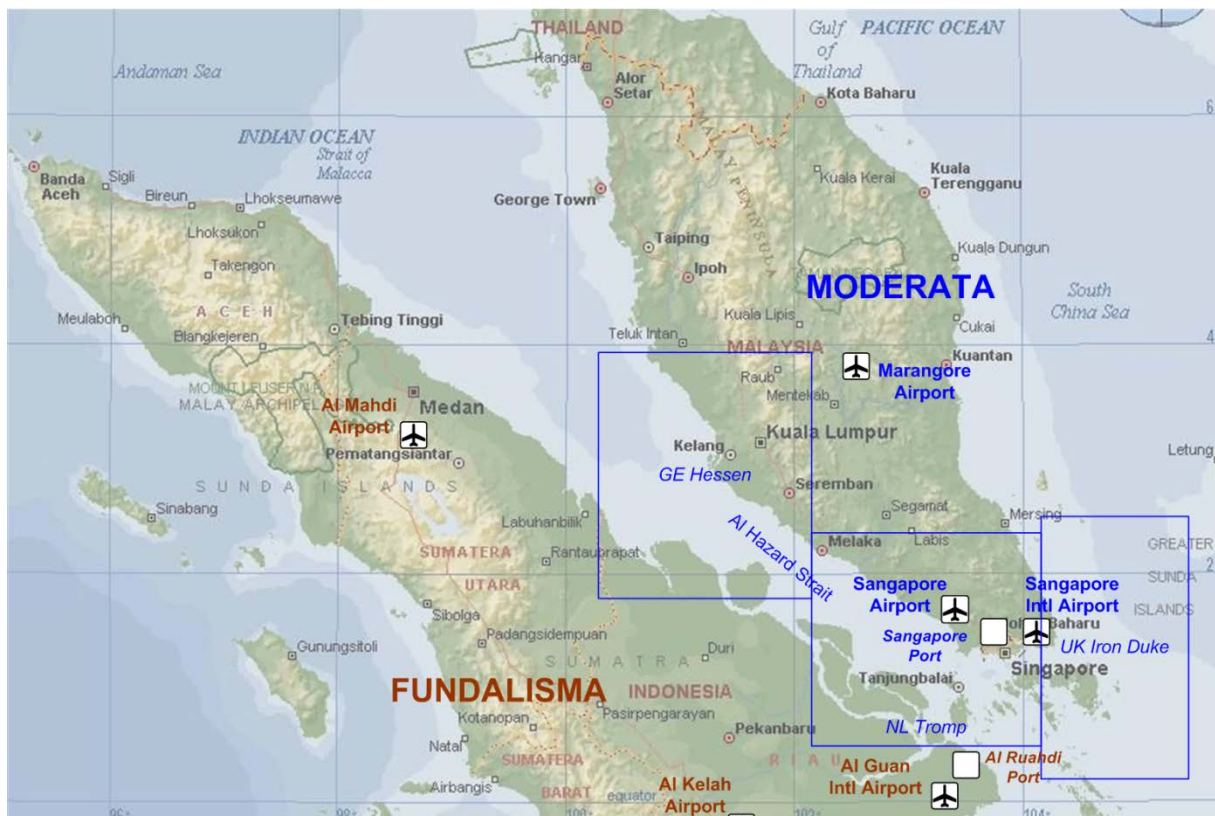


Figure 10. Operational Theater, which is modeled according to the sea strait between Indonesia and Malaysia.

Moderata itself is not able to effectively patrol the strait by a lack of seagoing ships. Moderata has only a small number of Coast Guard ships and several helicopters. With Moderata there is a good cooperation with the authorities (coastguard, police) and through the CTG can be warned for interception of ships or aircraft. The UN ships can also easily

territorial waters Moderata enter. The relationships with Fundalisma are less good. The territorial waters of Fundalisma are also closed to the UN ships. Fundalisma patrols by aircraft (fighters) and fast patrol boats (FPB) in the strait, both in territorial waters and in international waters. The patrol boats (FPB) and the hunters are armed with anti-ship missiles. There is a part of the Islamic world critical look at the efforts of the UN and the interpretation of it, in particular by Western forces. One purpose of Fundalisma is a conflict with UN ships to provoke that way the Islamic opinion and some of the world opinion on its side to get. Until a few months back also played much piracy in this area. With the growing presence of UN ships this decline, but there are still cases of piracy instead.

Appendix C. The Four SA Questions and Multiple-Choice Answers

	What is the identity of track #?
A	friend
B	Assumed friend
C	neutral
D	friend
E	hostile
F	unknown
	What is the behavior of track #?
A	en route or in transit (in airway or sea lane)
B	inbound to an airport or harbor
C	outbound to an airport or harbor
D	moving around in a bounded area
E	moving around unbounded
G	not enough data available
	What is the activity of track #?
A	transportation of goods or passengers for commercial purposes
B	localized commercial activity like fishing or oil winning
C	a private or recreational activity
D	a law-breaking activity e.g. smuggling.
E	military surveillance or reconnaissance
F	military provocation or attack
G	not enough data available
	What is the mission of track #?
A	conforms to a flight plan or sailing plan
B	conforms to a commercial area e.g. oil winning area or fishing grounds
C	conforms to a recreational area
D	conforms to a military CoA
E	does not conform clearly to any plan or area

Appendix D. Keyword Lists

Table 4. *List of keywords spotted for level 1 and level 2/3.*

Level 1	Level 2/3	
kilometer	aan boord halen	privé
boot	aangenomen vriendschappelijk	privé bezigheden
corvette	aanval	proeftocht
dalen	aanvallen	provocatie
dichtbij	afgebakend gebied	provoceer
onderzeeër	afwijken	recreatie
F16	assistentie	redden
fregat	binnen bereik	route
gevechtsvliegtuig	buiten bereik	smokkelen
helikopter	burger	te vroeg
hoog	commercieel	terrorisme
hoogte	doorreis	toerist
klimmen	escort	toezicht
knopen	evacuatie	transport
kust	goederentransport	transporteren
laag	illegaal	troepen transport
langzaam	intentie	varen
langzamer	lanceren	verdacht
links	logistieke ondersteuning	verkennen
luchthaven	lucht naar grond aanval	verkenning
mijl	luchtverdediging	verlaat
militair	neerstorten	vijandig
noord	neutraal	vissen
oost	olie winnen	vluchtplan
passagiersvliegtuig	onderscheppen	volgen
recht	onderwaterverdediging	volgens plan
rechts	onderweg	vriendschappelijk
richting	op de route	vuursteun
snel	ophalen	waarschuwing
stijgen	oppervlakteverdediging	
track	overtreding	
vaartuig	patrouille	
vissersboot	pers	
vliegtuig	personentransport	
west	piraterij	
zuid	plan	