Behavioural Marker Systems for evaluation of Norwegian merchant shipping Bridge Resource Management training

Candidate name: Tegegne Tefera

University of South-Eastern Norway

Faculty of Technology, Natural Sciences and Maritime Sciences

Masters Thesis

May 2019

Author Note

Correspondence concerning this thesis should be addressed to Tegegne Tefera; University of South-East Norway Faculty of Technology and Maritime Sciences; Email:012557@usn.no, tefera@mirkuz.com

Abstract

Human factors are fully or partially identified as the cause in 80 to 85% of all shipping incidents and accidents. Stakeholders invested considerable effort to ensure safe navigation and one of such efforts is Bridge Resource Management (BRM) training, an adaptation of the Crew Resource Management training, that was introduced to the aviation industry after a number of accidents in the 1970s.

International Maritime Organization mandated Bridge Resource Management training for all officers serving aboard a ship and substantial amount of resources are being spent; as a result, however, the instruments utilized to measure the effectiveness of the training are not sufficiently reliable.

Behavioural Marker Systems proved effective to evaluate Non-technical skills competency in aviation, operating theatres, nuclear facilities, and other safety-critical industries.

The aim of this research is to

1. Identify behavioral markers that are relevant for Norwegian merchant shipping bridge team Non-technical skills evaluation.
2. Establish their appropriateness by surveying for opinion of experts.
3. Calculate the relative importance of the identified behavioral markers.

The study has identified and shown that the Behavioural Marker Systems extracted from literature are relevant for Norwegian merchant shipping and calculated their relative importance.

Keywords: Bridge Resource Management, Maritime Resource Management, Behavioural Marker Systems, Non-technical skills, maritime

Acknowledgment

Leo Buscaglia said “change is the end result of all true learning”. Non-technical skills (NTS) and behavioral change has been a subject of interest for many industries to increase productivity and to stay competitive, however when it comes to safety-critical industries their impact is on life and environment. As an active sailor I am reminded of the many ways things can go wrong on the high seas, and learning how to prevent them was a solemn endeavour.

This journey was made enjoyable by the generosity of many to give their valuable time and patience. The highest gratitude goes to my supervisor Prof. Salman Nazir and co. supervisor Mr. (soon to be Dr.) Amit Sharma for their unwavering support and guidance.

Special thanks go to Norled AS human resources department and all the officers who took part in the survey. My gratitude goes to my shipmates who turned teachers, for their support and encouragement.

I would also like to thank the ALL Academy International AB for allowing me to attend the Maritime Resource Management (MRM) facilitator training.

I dedicate this work to my family. To my wife Lewtalem Begashaw and my son Mekuria T. Tefera, for your boundless love, support, and sacrifice.

T.T.

Tegegne Tefera

Contents

List of Tables

List of Figures

1 Introduction

"For safety is not a gadget but a state of mind."

— Eleanor Everet.

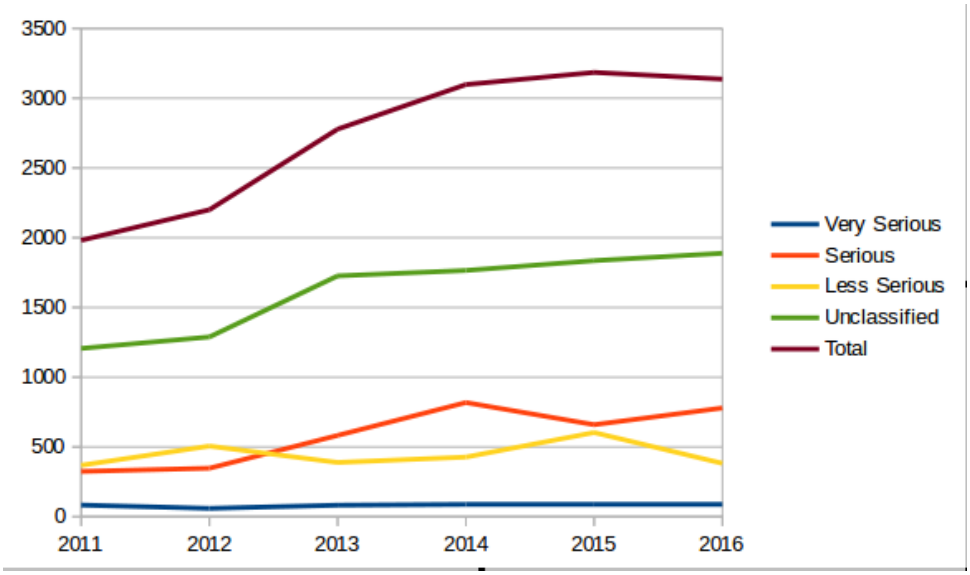
1.1 Research Background

Ships transport 80% of all global trade and 70% of its value. It is considered economical and environmentally friendly medium (United Nations Conference on Trade and Development, 2017). Yet, for long merchant shipping was considered a hazardous profession(Hansen, 2002; Nielsen & Roberts, 1999). The European Maritime Safety Agency (EMSA) reported, between 2011 and 2016, 18655 ships to have been involved in an accident and incidents. There were 253 ships lost, 16539 casualties, 5607 persons injuries, 600 fatalities and 869 investigations launched by EU member states(Emsa, 2017), out of which half were due navigational problems, such as collision and grounding. Chauvin, Lardjane, Morel, Clostermann, and Langard (2013) report that human factors are fully or partially identified as the cause in 80 to 85% of all incidents and accedents, costing substantially in terms of lives lost, cargo, and environment. In the last fifty years stakeholders (governments, shipowners, shipbuilders, clients, financiers among others) invested a considerable amount of effort on the design of better ships, education, training, and better working conditions to improve safe navigation (Bužančić Primorac & Parunov, 2016). One of this efforts is *Bridge Resource Management (BRM)* training which is the adaptation of the *Crew Resource Management (CRM)* training, which was introduced to the aviation industry after a number of accidents in 1970s who's, causes were related to assertiveness, leadership, fatigue, decision making and communication(Flin, O'Connor, & Mearns, 2002; O'Connor, 2011).

BRM's initial adaptation in the civilian maritime industry was to improve the interaction between the ship's captain and the pilot, who comes abroad to assist the ship in maneuvering through dangerous waters. It was later expanded and mandated to address the role of human factors on interaction, among the ships crew, the ships

systems and outside environment, in causing or escalating an emergency condition. Literature on the effectiveness of BRM training remains scarce compared to CRM despite the size of investment and the importance attached to it (O'Connor, 2011).

Figure 1. Number of marine casualties and incidents per severity



Note. adopted from Emsa (2017)

The International Maritime Organization (IMO) proposed and adopted a number of conventions to prevent accidents and reduce their consequence. One such conventions is the International convention on Standards of Training, Certification and Watch keeping for Seafarers (STCW) which was the first convention to establish basic qualification requirements for seafarers as its name indicates (IMO, 1995a). It was adopted in 1978 and came in to force on 28th of April 1984(Wu, Miwa, Shimamoto, & Uchida, 2015). In 1995 major revisions were made to clarify vague phrases, and to provide a mechanism for ease of administration and effective enforcement, and came in to force 1st of February 1997. Another comprehensive review was made in Manila, The Philippines, from 21st to 25th of June 2010 and adopted a number of amendments, among which BRM and Engine Resource Management (ERM) qualification were declared mandatory for bridge and engine room officers (Chauvin et al., 2013; Wu et al., 2015).

The search for an an independent definition of BRM in the available literature

refers to the CRM's broad term definitions. Salas, Rhodenizer, and Bowers (2000) define resource management training as a "family of instructional strategies" to improve *team work* by linking the training content (knowledge, skills, attitude) with well tested training tools such as simulators, videos, lectures. The STCW code table A-V/2 specifies the required knowledge, leadership and Team work skills, and the model course provided by IMO outlined a generic guidelines for training design, implementation and evaluations, which is not expected to be implemented literally but modified and improved to the needs of the trainees, organizational and national cultures(IMO, 2014).

1.2 Research Problem

An accurate assessment and feedback is essential to improve the effectiveness of a training(Arora et al., 2011). With respect to BRM assessment IMO provides generalized statements of evaluation criterion which is prone to subjective judgment and open to interpretation. In comparison to the civil-aviation industries, Non-technical skills (NTS) assessment criteria and competence standards in the maritime sector is in an early stage of development(M. Barnett et al., 2003).

The NTS definition this study is based up on is "the cognitive, social and personal resource skills that compliment technical skills, and that contribute to safe and efficient task performance"(Flin & O'Connor, 2017, p-11)

The basic concept of NTS, and corresponding BMS as identified by various researchers, is applicable for industries where human, internal and external system interactions are required to perform a critical task. However, the actual acquired and measurable skills and behaviours are specific to organizational, professional and national cultures. Therefore the training design, assessment tools development and deployment, needs to be within the context of the culture in question(M. Barnett et al., 2003). O'Connor (2011) points that, temptation to adapt a crew resource management developed for one domain to another may result in an ineffective training program.

The definition of BMS that underpins this study is given by Klampfer, Flin, and Helmreich (2001) as "observable, non-technical behaviours that contribute to superior

or substandard performance within a work environment".

1.3 Research Questions

The first part of the research is to derive BMS that can be utilised for the purpose of the research onjective by reviewing state of the art literature. Thus research question one is:

"What are the key BMS developed for safety critical industries?"

The literature review results will be presented to an expert group, to determine whether the behavioural markers identified by literature review are applicable to the Norwegian merchant shipping bridge team thus establishing content validity of the finding by using a five point *Likert like scale*. Hence the second research question is:

"Do the behavioural markers apply to the Norwegian merchant ships bridge team?"

The third part of the research is to determine the relative importance of each BMS, by employing the Analytical Hierarchy Process (AHP). The identified non technical skills and associated behavioural markers will be submitted to the expert group for a 9 scale pairwise comparison. Hence, the third research question to be answered is:

"What is the relative importance of the behavioural markers?"

The results of each research question is used as an input for the next research question.

1.4 Research approach and organization

In the Introduction section the research problem are identified and research questions are extracted. This will be followed by the Literature Review section which summarizes the existing state of the art literature on the subject and sets a theoretical framework to address the research problem. In the Research Methodology section the conceptual frame work, which methods and instruments are to be used to answer the research questions, and their rationals are explained.

The Research results and Analysis section discusses the main findings of the research, and in Discussions section the significance of the findings with respect to the

research problem and research questions will be discussed. The Conclusion section discusses the research contribution and closes with recommendations for further research. Table 1 shows the research approach and organization.

Research questions, Methods and Objectives

Research questions	Methods	Objectives
Q1. *What are the key Behavioural Marker Systems developed for safety critical industries?*	Literature Review	Discuss and identify BMS
Q2. *Do the behavioural markers apply to the Norwegian merchant ships bridge team?*	Expert opinion using Likert type scale	Verify applicability to the Norwegian merchant ships bridge team.
Q3. *What is the relative importance of the behavioural markers?*	Using Analytical Hierarchy process	To construct BMS model that can be used for evaluation of BRM training.

Table 1
Research process

2 Literature Review

2.1 Scope of Literature review

A review of the state of the art literature carried on this chapter. *Google scholar, Oria the Norwegian academic libraries search engine,* and *Scopus Citation and Abstract database* were used to locate previous research materials. Combinations of the following keywords and phrases were used to locate relevant literature. *Bridge Resource Management*, *Maritime Resource Management*, *Behavioural Marker Systems*, *Non-technical skills*, *maritime*, *Norway.*

2.2 Literature search

A search using the *Behavioural Marker Systems* and *maritime* keyword combination produced 23 documents out of which the majority are from the United Kingdom (11) whereas Australia, Norway, and United States produced three each. *Non-technical skills*, *Behavioural Marker Systems* and *maritime* keyword combination produced seven documents that overlap with preceding search and mostly of offshore installations that have little in common with the shipboard activities. In recent years training institutions combined the Bridge and Engine resource management courses, which run in parallel and coordinated for real shipboard experience, and renamed it as *Maritime Resource Management (MRM)*(ALL Academy, 2019). A search for the same term produced one research paper, which was meant to frame a research agenda on the topic (M. Barnett et al., 2003).

To find if there is literature on development of BMS in Norway the *Behavioural Marker Systems*, *Norway*, keyword combination search produced one document, a BMS development research of Nurse Anaesthetists' Non-Technical Skills-Norway (NANTS-no)(Flynn, Sandaker, & Ballangrud, 2017) which is the adaption of Anesthetists' Nontechnical Skills (ANTS)(Patey et al., 2005). Øvergård, Sorensen, Hontvedt, and Nazir (2017) noted the absence of BMS for a case studies of simulator based BRM trainee evaluation at the University of South-East Norway (USN). There was no document found by literature search where BMS are developed or used for

Norwegian merchant shipping NTS evaluation.

The literature review will use the summary of the proceedings of the *Behavioural Markers Workshop*(Klampfer et al., 2001) and literature on three prominent projects; the NASA and the University of Texas Human Factors Research Project (NASA/UT), the NOTECHS, and the Group Interaction in High Risk Environments (GIHRE) as a foundation. Regarding BMS for maritime sector the work done by M. Barnett et al. (2003); P. M. Barnett et al. (2006) and Saeed, Wall, Roberts, Riahi, and Bury (2017) will underpin the literature review and methodology. Relevant references given in the above literature will be consulted. On the Scopus database weekly alert was set, between week 10 and 30, 2018 to follow any new developments on the subject.

Existing theories of training, adult education, and evaluation methodologies will be discussed, to set the theoretical framework for the study.

2.3 Learning vs Education

2.3.1 Learning. Learning and how it takes place intrigued thinkers for generations. It is a fundamental part of human nature, which without, human development is inconceivable. (Jarvis, 2012b, pp.200) contends that all learning is experiential, that is affected not only by our immediate but past experiences and the social order around us. It is a process that is self or outside agent initiated and ever transforms the person physically, mentally and emotively.

As the understanding of learning expands, involving various outlooks and disciplines the traditional definition of learning as *acquisitions of knowledge and skills* is no longer adequate, and as a result there is no common agreed definition of the term. (Illeris,2003,p.54-57;Jarvis.

Jarvis summarized a number of competing and overlapping learning theories, such as *behaviourist*, *cognitive humanest* and *social*. Learning, with the adult population perspective, coined as *andragogy* was further elaborated by Knowles (1978), which takes in to account the *self concept*, *experience*, *readiness*, *orientation*, and *motivation* to learn, in contrast to *pedagogy* whose roots are in child education, though this distinction

is given less weight in latter studies(Jarvis, 2012a).

2.3.2 Education. In contrast to learning, education is prescribed learning, designed to create successors to a community. It is restricted by criterion set by providers, as such requiring assistance in building self-sustaining interest. The expectation that, at formal education institutions, an individual would be equipped with enough Knowledge and skills to serve for a lifetime is no longer holds. The rapid economic growth, increased competition, technological innovation, and globalization that followed the second world war necessitated the acquisition of new skills and updating existing once, bringing about the notion of *life long learning*. Consequently, the focus of education is increasingly becoming, to prepare the learner to be a lifelong learner in addition to imparting Knowledge and skills. (Jarvis, 2012a; Masadeh, 2003).

Among human resource development professionals, there is an overlap in the usage of the terms *Learning*, *Education*. *Development*, *Training* (Garavan, 1997; Masadeh, 2003). The author finds for this thesis the summery of Masadeh (2003) most appropriate. He defined training as teaching facts and a hands-on opportunity to learn by doing, while development is an overarching concept and long-term road-map to enhance an individuals ability through learning and experience beyond the *initial education*.

2.4 Transfer

Transfer is defined as, a degree to which learning of a response to one task and context influences a response to another task and context. For a transfer to have occurred the Knowledge, skills, and attitudes gained must be generalized and applied to another task and maintained over a period. The degree of effectiveness of training is expressed by the degree of *positive transfer* occurred. Where the training is performed called *training environment*, the actual task is performed at a *transfer environment*(Baldwin & Ford, 1988; Blume, Ford, Baldwin, & Huang, 2010). Therefore, transfer can be said to have taken place, if a skill learned at a simulated navigation training, improves the performance of actual navigation of a ship.

The significance of spatial, temporal, and context proximity of a training and transfer environment for positive transfer is denoted by *near* and *far transfer*. Blume(1988) explains the degree of similarity of a training environment to a transfer environment creates a similar stimuli-response relationship. Furthermore, a transfer will occur where "aims, methods, and approaches" employed on a task at a training environment are used on a similar task at transfer environment. When transfer occurs across tasks with a similar level of complexity it is called *Lateral transfer* while *vertical transfer* is said to have taken place if a skill acquired helps in acquiring skill with higher complexity. Moreover *Soft-skills* (interpersonal and intrapersonal) face more difficulty to transfer than *hard-skills*(technical)(Laker & Powell, 2011).

2.5 Training design

Studies have shown the choice of training design affects training effectiveness, therefore transfer significantly(Blume et al., 2010; Ritzmann, Hagemann, & Kluge, 2014). The generic *Analysis, Design, Develop, Implement, Evaluate (ADDIE)* training design process is a systematic approach to identify training needs, design and develop training instructions and resources, implement the program and evaluate the outcome to measure the effectiveness. (Aldoobie, 2015; Allen, 2006).

2.6 CRM Training

The root of CRM lies on the research NASA conducted on the cause of air accidents that culminated in the workshop of 1979, that identified, "interpersonal communications, decision making, and leadership" or in other word *human error* as the primary cause for the mishaps. The first generation CRM training's were modeled after techniques developed to enhance managerial effectiveness for corporations (Salas et al., 2000) and "grounded in social, cognitive and organizational psychology as well as in human factors research"(R. L. Helmreich, 1997). Intensive seminars were delivered that included diagnosing once managerial style and correcting deficiencies such as "lack of assertiveness by juniors and authoritarian behavior by captains."

The second generation CRM training's were the result of another NASA workshop

conducted in 1986 that utilized prior experience to draw the attention from the individual pilot to the flight operations, which brought forward concepts such as "team building, briefing strategies, situation awareness, and stress management." The 3rd generation CRM training of the early 1990s brought in to focus, issues such as organizational culture and extended the training beyond the cockpit crew to flight attendants, maintenance, and dispatch personnel. The fourth generation training were introduced by the US Federal Aviation Administration with the introduction of Advanced Qualification Program (AQP) which allowed the airlines to adopt the system to their needs with the condition that (a) "complete detailed analysis of training requirements of each aircraft", (b) employ formal full mission evaluation, (c) training requirement for those responsible with certifying(R. L. Helmreich, Merritt, & Wilhelm, 1999).

2.6.1 Training design challenges. Poor generalizability of CRM to export across cultural divides, lack of validation of training programs, rejection of the training by trainees, and the decay of training over time, are raised as some of the challenges faced(R. L. Helmreich, Merritt, & Wilhelm, 1999; Salas et al., 2000).

"Culture is the collective programming of the mind that distinguishes the members of one group or category of people from others" (Hofstede, 2011). Hofstede's cultural dimensions are relevant in determining the design and effectiveness of resource management training. In *high power distance* cultures where respect for senior officers is strictly enforced the challenge would be to train senior officers to accept suggestions from their Juniors and to encourage junior officers to be assertive, however, these same cultures are more collective; therefore, collaboration skills are easily understood and accepted. In low power distance cultures, on the other hand, both senior and junior officers would have no problem with assertive behavior but would have difficulty to relate to teamwork as low power cultures, as they tend to be more individualistic. The uncertainty avoidance dimension that refers to the degree of tolerance to ambiguity also influences the design of resource management training. Those from cultures with a higher tolerance for ambiguity may find CRM training unnecessary, while those from

cultures that do not tolerate uncertainty welcome it. These properties require *culture* to be taken into consideration in the design, implementation, and evaluation of CRM training (R. L. Helmreich, Merritt, & Wilhelm, 1999).

2.6.2 From error avoidance to Error Management. The initial drive behind CRM was human error and the need to avoid it, but how to achieve it was not very well defined(R. L. Helmreich, Merritt, & Wilhelm, 1999). The fifth generation of CRM was based on *error management* theory of Reason (2000) which gives generality to the training for application across cultures.

Reason gives two perspectives on human error management. The *person approach* and the *system approach*. Person approach treats errors as violations by individuals due to "aberrant mental process" such as lack of motivation, negligence, loss of attention, and the recommended remedy, therefore, become more procedures and "reducing unwanted variability in human behavior" by treatment of the senses through threat or shame (Reason, 2000). Experience shows that this approach is a dominant tradition in the maritime industry.

However, the premise of *system approach* to error management is, "humans are fallible, and errors are inevitable". Therefore incidents are treated as a source of valuable information to build barriers and safeguards, and the error management program targets the institution as a whole, which includes "the person, the team, the task, and the workplace". For this approach to work organizations must forgo the blame culture and adopt a "none-punitive" policy for unintentional errors to encourage reporting. In such management, CRM can be viewed as an error countermeasure for human performance limitations, by equipping with three lines of defense skills. That is *avoid incidents*, *tolerance(trap) errors*, and *containing the after effects of errors*(R. L. Helmreich, Merritt, & Wilhelm, 1999; Reason, 2000). The result was the development of the *Thereat and Error Management (TEM)* that was derived from the *Line Operations Safety Audit (LOSA)* system, became the basis for subsequent CRM training. The LOSA methodology was utilized on normal flights, to collect none-jeopardy data of errors and their management systematically, and included *BMS*

to asses the *Non-technical skills* addressed(R. L. Helmreich, Klinect, & Wilhelm, 1999; Klampfer et al., 2001).

2.7 Training Evaluation

Evaluation is defined as "a systematic collection of descriptive and judgmental information" on training, necessary to make an informed decision regarding an implementation, modification or value of training (Ritzmann et al., 2014). Often, *training evaluation* and *training effectiveness* are used interchangeably, however they are different constructs. While training evaluation is about measuring the degree of success in achieving training goal (transfer), training effectiveness measures the effectiveness or lack thereof of training characteristics such as training design, environment, personnel, as such nested within training evaluation. (Alvarez, Salas, & Garofano, 2004).

The four levels Kirkpatrick's hierarchical model is an important framework to categorize training outcomes.

Level 1: Reactions. Reactions are measures of enjoyment of training, and its advantage includes time and cost-effectiveness and relatively easy to implement(Twitchell, Holton, & Trott, 2000). Reactions is further decomposed as *affective* reaction (enjoyment of training), *utility* reaction (usefulness of training) and perceived difficulty(Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997). In a decomposed construct *affective reaction* correlates weakly with learning and behavior transfer, however *utility reaction* relates strongly, therefore, a good predictor of transfer (Warr, Allan, & Birdi, 1999). *Perceived difficulty* "correlates significantly with self-reported competence, Knowledge, self-reported use of training on the job and perceived value of training." (Ritzmann et al., 2014). When Affective and utility reactions and perceived difficulty are measured, reaction data can provide valuable information in training evaluation. Reactions treated as a combined construct show very small correlation to *learning* and *behaviour* though it predicted post-training *declarative* and *procedural* knowledge.

Level 2: Learning. Common objective of training are acquisition of *knowledge* (cognitive or declarative), *skill*(procedural, automaticity) and a change in value and attitude (affective, self-efficacy, motivation)(Kraiger, Ford, & Salas, 1993; Ritzmann et al., 2014; Warr et al., 1999). *Learning* is measured as an outcome of training than as a change in behavior. Questionnaires, exercises and work samples can be used to measure learning and attitude. *Declarative Knowledge* can be assessed by tests on the training content and attitude questionnaires asses attitude towards the attitude object. Comprehensive evaluation needs data on multiple levels, yet most evaluations are restricted at reaction level to save time and cost. (Ritzmann et al., 2014; Twitchell et al., 2000). However, this was criticized as there are results that show enjoyment of a training course does not necessarily correlate to learning or transfer of behavior (Alliger et al., 1997; Ritzmann et al., 2014).

Ritzmann et al. (2014) developed a training evaluation method *The Training Evaluation Inventory (TEI)* by combining Kirkpatrick's three component variables of level one , *enjoyment, perceived usefulness, perceived difficulty*, level 2 variables, *subjective knowledge, attitude* with design dimension of the *First principles of instruction* Merrill (2002), but they pointed out that "the TEI is not suitable to answer questions such as whether or not trainees are able to apply a learned technique in practice".

D. Kirkpatrick (2007) points out that level two is important to improve training content as it measures the "effectiveness of the instructor and instruction", by analyzing the change in the answers. Without cognitive, affective or skill related change, on-the-job behavior change and organizational results cannot be achieved.

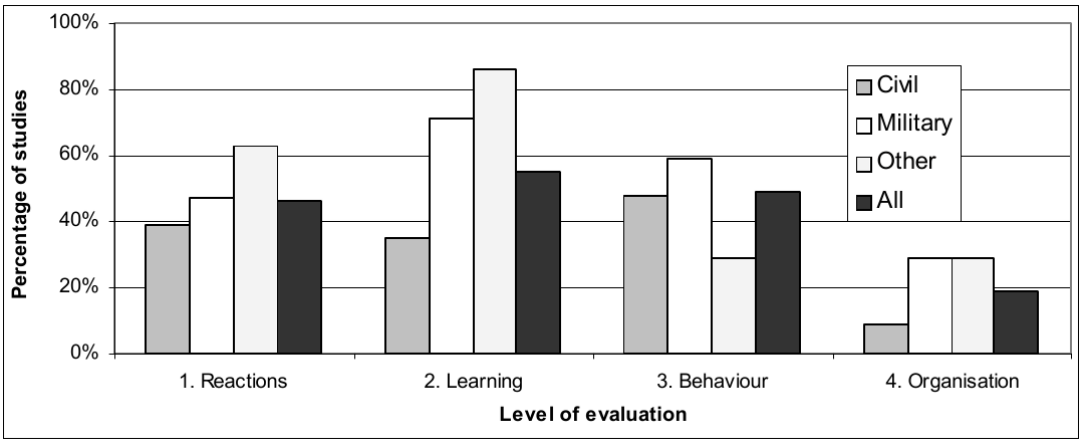
Level 3: Behavioral. Behavioural level evaluates whether the Knowledge acquired did transfer to change of behavior on the job and to what extent.
D. Kirkpatrick (2007) points that this is the more complicated evaluation compared to *reactions* and *learning* levels . Though one may acquire Knowledge, lack of opportunity to demonstrate, the uncertainty of when the change of behavior occurs, attitude towards the Knowledge itself makes measuring behavior change a challenge. Behavior

level evaluation is done in CRM training using behavioral marker systems(O'Connor, 2002). J. D. Kirkpatrick and Kirkpatrick (2016)(p. 84-86) describe level 3 as a "continuous performance monitoring and improvement mechanism", and the evaluation is done as performance measure on identified few *critical behaviours*, which are *observable* and whose quantity(number of times) and quality(degree) is *measurable*.

Level 4; Results. Results level measures training outcome to the organization in terms of cost saving, increased productivity or reduced accidents(Ritzmann et al., 2014). D. Kirkpatrick (2007) point that evaluation process runs opposite the design process. While designing a training program, what results the company hopes to achieve need to be established first. The behaviors needed to achieve those results, the Knowledge, skills, and attitudes needed for the emergence of the behavior, the type and delivery of training that is best suited to deliver those skills are considered.

Availability of data on multiple levels allows a more comprehensive evaluation of training outcome. However, organizations are restricted at reactions level due to time and cost constraint(Ritzmann et al., 2014). O'Connor (2002) discovered that out of 48

Figure 2. Percentage of studies Carrying out CRM evaluations



Note. Adopted from (O'Connor, 2002)

studies considered CRM training's majority of training evaluations are done at learning level(knowledge and/or attitude). However data in civilian aviation as shown in figure2, most evaluations are done at behaviour level, except two that evaluated at organization level, and the technique that is used widely to assess CRM skills is called *behavioural*

markers systems.

2.8 Behavioral marker systems

Behavioural Marker Systems and Non-technical skills taxonomy are a CRM performance and training evaluation method for *safety-critical* industries(Klampfer et al., 2001). The term Safety-critical industries is used for industries where the consequence of failure or malfunction may result in loss of life, serious injury, or serious environmental damage. Industries such as commercial aviation, shipping, nuclear plant are given as examples(Amalberti, 2001; Lowe, Hayward, & Branford, 2016).

BMS are structured hierarchically from a broad category to constituent elements, as a taxonomy of skills and exemplar behaviours(Thomas, 2017), and terms *Behavioural Marker Systems*, *Non-technical skills*, and *Interpersonal Behaviors* are used interchangeably(Klampfer et al., 2001). Generic structure of BMS is shown in figure 3.

Behaviour is an observable action(Schrader & Lawless, 2004) and the presence of skill or Knowledge can be observed in behavior; therefore, BMS are, observable behaviors that indicate the presence of skill or Knowledge. BMS developed as a research tool are more complex and differ in detail level and scope from that of meant to be used as a base for training or assessment(Flin & O'Connor, 2017).

The techniques used to develop BMS are given as follows.

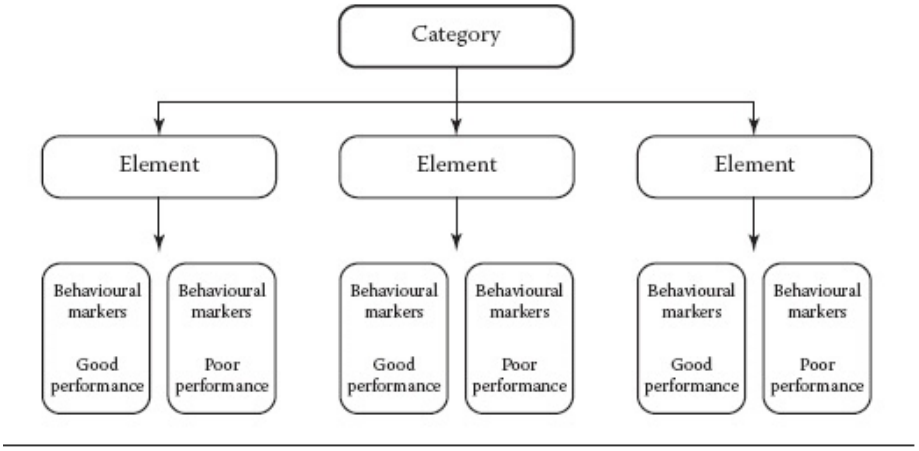
- review of scientific literature,
- event-based analysis, such as incidents and accidents report analyses,
- interviews, surveys, focus groups, ethnographic studies,
- cognitive task analysis,
- direct or remote observation of conducting task (simulator, video)(Klampfer et al., 2001; Thomas, 2017)

A good BMS have a causal relationship to performance; should be described in a clear, operational environment specific language, and should have minimum overlap

between categories and elements. They are objectively observable in normal operations and training, and describable in *word-picture*, and do not focus on *attitude* or *personality trait*(Klampfer et al., 2001; Thomas, 2017).

The development of BMS should be followed by assessment of validity (the degree of actual assessment to which the BMS claims to assess)and reliability(constancy of results as in *Inter-rater reliability*) as well as *sensitivity* to performance level, prior to using it for training assessment(Thomas, 2017, p.105). *Transparency* to the observed and *usability* to the trainers and observers are important attributes of BMS (Klampfer et al., 2001).

Figure 3. Generic structure of a behavioural marker system.



Note. adopted from (Thomas, 2017, p.101)

2.8.1 NASA/UT project. The original research and the first set of behavioral markers taxonomy for the aerospace crew was done by the NASA and the University of Texas Human Factors Research Project (NASA/UT) in the late 80s to evaluate the effectiveness of CRM training. Before the NASA/UT project, crew performance assessment was made by Line Oriented Flight Training (LOFT) which was an observation of the entire flight from initial briefing to engine shutdown, or full mission simulation, which uncovers proficiency issues but not CRM training performance (Klinect, Murray, Merritt, & Helmreich, 2003). Following the NASA/UT project many airlines developed their own behavioural markers systems (Crayton, Hackworth, Roberts, & King, 2001; Klampfer et al., 2001).

By the mid and late 1990s knowledge gained on *human error* and *systems thinking* brought the concept of *Thereat and Error Management (TEM)* in to the picture, which provided an empirical criteria to correlate with BMS. The term *Line Operations Safety Audit* was used to reflect the shift of focus from a person-centered approach to system-centered approach. The audit provides a snapshot of safety performance, strengths, and weakness, in a formal process that allows trained observers to record their observation of normal operational activities in a non-jeopardy assessment(Crayton et al., 2001; Klampfer et al., 2001; Klinect et al., 2003).

However, researchers discovered a high degree of variation in CRM performance rating, within the same airline, flying different type of aircraft. Across airlines, a significant performance difference was registered between the markers, and research failed to establish the case of the difference, whether it is due to organizational difference or due to emphasis on different aspects of CRM skills. Moreover, evaluation of the same crew can vary considerably by different raters, which led to the realization that, before the validity of CRM assessment can be properly measured, it is important to standardize the raters (Flin et al., 2002).

Table 2
LOSA BMS

Planning	Execution	Review/modify Plans
• Briefing	• Monitor/ cross-check	
• Contingency *management*	• Workload *management*	• Evaluation of *plans*
• Workload *assignment*	• vigelance	• Inquiry
• Plans *stated*	• Automation *managmenet*	• Assertiveness

Note. Adobted from (Dietrich & Chhildress, 2017)

2.8.2 The NOTECHS project. Non-Technical Skills assessment system (NOTECHS) is a result of a project initiated by four European Civil Aviation authorities in collaboration with the University of Aberdeen (UK) to fulfill the European Joint Aviation Requirements (JAR) requirements which states

> the flight crew must be assessed on their CRM skills under a methodology acceptable to the Authority and published in the Operations Manual. The purpose of such an assessment is to: Provide feedback to the individual and serve to identify retraining; and be used to improve the CRM training system (Flin et al., 2003; O'Connor et al., 2002).

The development of the behavioral markers for NOTECHS consisted of three phases. These are (a) review of existing pilot proficiency evaluation systems, to identify common categories and elements; (b) literature review of relevant research findings relating to key none-technical skills categories identified in existing systems; and (c) extended discussion with subject matter experts (experienced non-technical skills evaluating pilots). To make the system applicable for all airlines and for all languages across Europe the following design criteria were set.

- maximum exclusivity of categories and elements,
- a rule of parsimony,
- Terminologies used in everyday language should be used,
- Skills should be
    - inferrable in the case of cognitive skills
    - observable in case of social skills(Flin et al., 2003).

NOTECHS is structured in four main categories: *Co-operation, Leadership and Managerial, Situation Awareness,* and *Decision Making*, that can be placed under two broader skill sets: *social and cognitive.* The main categories are further divided into skill elements with exemplar behaviours(Flin et al., 2003).

Validation of NOTECHS. Subsequently, the Joint Aviation Requirements: Translation and Elaboration of Legislation (JARTEL) project was launched, by a consortium of European aviation industry and Universities, with the objective validating the methodology proposed by NOTECHS. The following tasks among others were designed and carried out.

Table 3
NOTECHS BMS

Co-operation	Leadership and Managerial Mkill	Situational Awareness	Decision Making
• Team building maintaining	• Use of authority assertiveness	• Awareness of aircraft systems	• Problem definition and diagnosis
• Considering others	• Providing, maintaining standards	• Awareness of external environment	• Option generation
• Supporting others	• Planning co-ordination	• Awareness of time	• Risk assessment/ option selection
• Conflict solving	• Workload management		• Outcome review

Note. adopted from (Flin et al., 2003)

• Review its claimed cultural robustness as it would be implemented across cultural differences and cultural dimensions;

• developing experiment protocol that defined scripts and scenario videos to be shown at the classroom validation to the instructor;

• production of videos, based on the protocol and scripts developed on phase two, and evaluate and analyze;

• additional confirmation in the form of operational validation of NOTECHS; and

• production of guidelines for operational implementation for airlines(JARTEL_Consortium, 2002).

The project concluded that NOTECHS demonstrates acceptable sensitivity for proficiency variation, to a great extent culturally robust, and found to be practicable by the participating pilots, and can be used to consistently evaluate NTS skills.

2.8.3 The GIHRE project. The Group Interaction in High Risk Environments (GIHRE) was a (multidisciplinary (cognitive, experimental, social, and organizational psychology as well as linguists and psycho-linguists) effort from united states, Switzerland and Germany that run between 1999 and 2004 by collegium of the Gottlieb Daimler and Karl Benz Foundation to study the behaviour of professional teams in safety-critical work environments. Seven groups (Threat/Error, Behavioral Markers, Process Control, Linguistic Factors, Co-ordination, Language Processing, and Micro-structure) investigated four safety-critical work place teams as a whole (airline

cockpit, nuclear power plant control room, an intensive care unit of a hospital and an operating room) , to devise practical suggestions for enhancement of performance in the said work environments. The *Behavioural Markers Group* conducted three studies which are,

- the impact of task load on team performance,
- the stability of CRM performance in different situations including persons character traits
- comparison of the LOSA and NOTECHS behavioural systems(Dietrich & Chhildress, 2017, pp.9,11,34-35).

Three variables were employed for measurement.

- *CRM performance* measured using the NOTECHS and LOSA BMS systems.
- *Risk index*, an aggregate measure of substandard performance of CRM as registered by BMS systems, which indicates the degree to which a crew behaviour compromises flight safety.
- *Subjective work load* which is defined as "an effort invested by a human into task performance" which has six sub-scales, *physical demands*, *mental demands*, *temporal demands*, *performance*, *effort* and *frustration* (Dietrich & Chhildress, 2017, p. 39-40)

A quasi-experiment design on an airbus 320 simulator, where three scenarios and different levels of task load were analyzed on 46 crews out of 81 volunteered.A camera monitored behind the crew, the instrumentation data, and the questionnaire filled by pilots and flight instructors was used for the analysis. Three external observers with formal training in LOSA and NOTECHS were included in the study while subject matter experts reviewed *poor* and *very poor* rating to prevent any systemic or rater bias. *Inter-rater* agreement was measured and calibrated to enhance the quality of judgment. The study observed that though there is an important distinction and

application purpose between the LOSA and NOTECHS (while the former is designed for safety audit on normal flights and research purpose the latter is a tool to evaluate the performance of individual pilots for training and qualification purposes) they contain similar constructs as it can be seen from table 2 and table 4. Independent evaluation of the same scenarios using LOSA and NOTECHS BMS similar constructs showed a moderate to high correlation(Dietrich & Chhildress, 2017, pp.34-48).

2.8.4 BMS in shipping. Though it is reported one third decrease in accedents, per ship year at Maersk shipping four years after the introduction of CRM training(Flin et al., 2002) researches conducted in to the maritime CRM training's did not find evidence of effectiveness, likely due to the adaptation of the training "as is" from the aviation than adopting it to the need of the maritime industry through a systematic review (P. M. Barnett et al., 2006; O'Connor, 2011). While there is a more comprehensive research to develop BMS for the navy (F. P. Da Conceição, Basso, Lopes, & Dahlman, 2017; O'Connor & Max Long, 2011), literature on the merchant shipping BMS development is scarce. At Warsash maritime academy BMS adopted from a Rail Safety and Standards Board (RSSB) were validated in 60 simulator exercises over four months period. They concluded "the overwhelming majority of the behavioural markers appear to be both relevant and observable" with few exceptions such as watch handover.(Devitt, Holford, Pantaleev, & Sharma, 2012).

Warsash maritime academy have further taken *reflective practitioner* approach in its maritime CRM training with debriefing session after the training and a follow up three months later. The followup helps to identify whether the organization culture is conducive to implement the newly acquired non-technical skills (P. M. Barnett et al., 2006).

Saeed et al. (2017) at Liverpool John Moores University, School of Engineering, Technology and Maritime Operations, used interview method, with 12 senior deck officers to develop a taxonomy of NTS and behavioural markers with five performance level. The identified behavioural markers relative importance was calculated using the Analytical Hierarchy Process (AHP) method.

Table 4
BMS

Social skills	Cognetive Skills
Co-operation	**Situational awareness**
• Open communication	• Situation assessment
• Consideration for others	• Risk assessment
• Team working	
Leadership and Managerial Skill	**Decision Making**
• Situational leadership	• Problem Diagnosis
• Assertiveness	• Option Generation
• Planning and coordinating	• Option selection

Note. adopted from(M. Barnett et al., 2003; Flin et al., 2003)

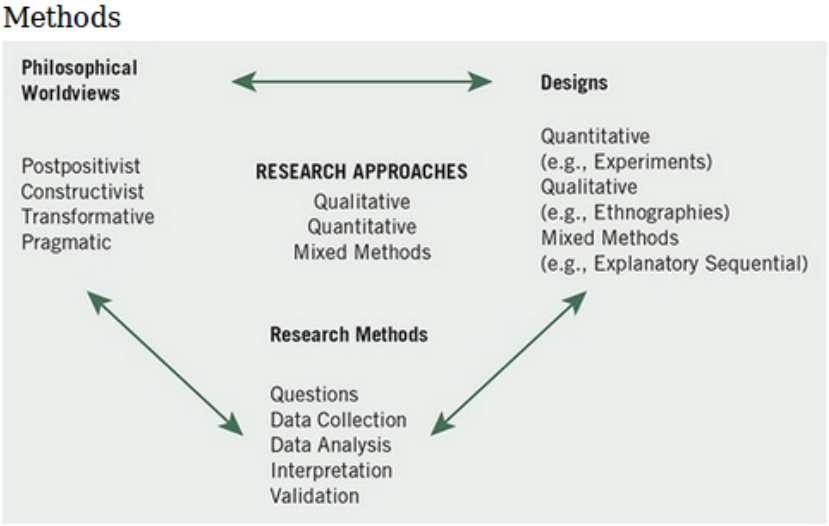
2.8.5 BMS development in Norway. At the time of writing there was found one BMS adaptation and validation effort, NANTS-no, performed from an approved translation of ANTS in a quasi-experimental, pre-test post-test, simulation-based study(Flynn et al., 2017).

2.9 Summary.

The literature review set-out to investigate the state of the art literature in training and evaluation particularly in *safety-critical industries*, to develop an evaluation instrument for BRM training for Norwegian merchant shipping. The development of BMS was the result of the need for a more practical evaluation tool for CRM training. The review identified BMS development efforts and practices, and found no BMS development effort for the merchant shipping sector in Norway.

3 Research Methodology

Figure 4. A research framework



Note. adopted from Creswell and Creswell (2017)

3.1 Research Framework

A *research framework* is an interconnection between a research *philosophical worldview*, *research design*, and *research methods*.

3.1.1 Research world view. Research worldview or paradigm as defined by Guba (1990) is "a basic set of beliefs that guide an action" and points to the philosophical orientation of the researcher in approaching the research problem and is dependent as much on the discipline and the research community as on the researcher's experience. There are four worldviews defined by literature; post-positivism, constructivism, transformative, and pragmatism(Creswell & Creswell, 2017, p. 35). *Post-positivist* worldview got its name due to its challenge to the notion of *absolute truth* of positivism. By stating that we can not be absolutely positive about our claim of knowledge, about human behavior, or action, however, maintaining the traditional empirical research approach, where research begins with a theory then collect data to support or refute, thereby refining incrementally or abandoning theories depending on their merit. The *constructivist* worldview, on the other hand, attempts to understand

meanings humans construct as they interact with the world. The production of meanings, therefore, is social and the cultural context of participants and the researchers own experiences and background influences the interpretation of the observation. The *transformative* worldview holds the view that the post-positivist approach did not heed to the problems of disadvantaged communities. Accordingly, it intertwines a research inquiry with a socially transformative agenda, taking cue from one of the challenges, involving the participants as collaborators instead of marginalizing them from the process of observation, and producing an action agenda to transform their lives. The *pragmatic* worldview does not adhere to any discipline; it instead concerns itself with the research problem, solution, and application, "truth is what works at the time" and uses a pluralistic approach to gain knowledge about the problem(Creswell & Creswell, 2017, p.36-42).

3.1.2 Research design. A research design is about a choice of one of the three, *quantitative, qualitative*, and *mixed methods* design and includs procedures and strategies to direct the inquiry. Quantitative research attempts to determine if a theory explains/predicts a phenomenon of interest by measuring the relationship between an independent and dependent variable in a true experiment where factors that affect the research interest are controlled, or quasi-experiments are used where control of all the factors is not practical(Yilmaz, 2013).

A qualitative research design is deemed challenging to define and used as an overarching category of many paradigms and approaches. The component methods of qualitative design such as ethnography, grounded theory, case studies are better defined. Yilmaz (2013) incorporates the essential elements proposed by other researchers and defined qualitative method design as, "an emergent, inductive, interpretive and a naturalistic approach to the study of people, cases, phenomena, social situations and processes in their natural settings". The goal is to unveil and describe the significance of the experience. *Mixed methods* design combines quantitative and qualitative methods as the name implies and the early rationale behind the method was that all methods have bias, and the combination of the two approaches would improve the quality of the

research by one validating and converging, explaining and exploring sequentially to provide better instruments and database to the other(Yilmaz, 2013, p. 45).

3.1.3 Research method. The third component of the research framework is the research method which deals with the data collection analysis and interpretation. Depending on the research subject and research method design, the data analysis can be numeric or non-numeric, and the questioning method close-ended or open-ended or both(Yilmaz, 2013, p.47). Since the development of Behavioral markers will be qualitative as well as quantitative, the research design approach this thesis will follow will be a pragmatic worldview and a mixed research method.

3.2 Research process.

3.2.1 Pilot survey. The resulting BMS will be presented to three actively sailing bridge officers to test the face validity, and ease of understanding on a simple yes or no scale. The validity can be established from the percentage of agreement on each category (Thorn & Deitz, 1989).

$$P_{yi} = \frac{N_{yi}}{N} \tag{1}$$

where:

P_{yi} is a percentage of yes for item i

N Number of experts

N_{yi} Number of judges assign yes for item i

3.2.2 Content Validity. On the second part of the study, the extracted behavioral markers will be presented to the Norwegian expert group(Captains and Chief officers), to collect their opinion on their applicability to the Norwegian context, in the form of a five-point *Likert type scale* survey. The *consensus* and *content validity ratio(CVR)* of the survey will be calculated thereby answering research question number two.

Consensus towards an issue is a function of shared group "feeling" captured by *likert like scale*. The subjective but informed opinions of experts are converted to

ordinal value to represent the extent of agreement or disagreement to the statement of interest(Tastle & Wierman, 2007). Consensus is given by the formula

$$Cns(X) = 1 + \sum_{i=1}^{n} p_i \log_2 \left(1 - \frac{|X_i - \mu_x|}{d_x}\right) \tag{2}$$

where:

μ_x is the mean of X and d_x is the width of X given by $d_x = X_{max} - X_{min}$. Tastle and Wierman (2007) further note that the following rules must be satisfied for a measure to be considered a viable solution to Likert type scale consensus problem. These rules are

1. For a given even number of participants, if $n/2$ number of participants select *agree* and $n/2$ select *disagree* categories, then the group has *no* or 0 consensus.

2. If all participants select the same category then there is 100% consensus and returns a value of 1.

3. If the mix of $n/2 + 1$ participants selects any one category the degree of consensus must fall between 0 and 1.

Content Validity is the measure of domain relevance (Salkind, 2010)of each element/category in the context of BRM and will be calculated by using the (Lawshe, 1975) equation for Content Validity Ratio (CVR).

$$CVR = \frac{n_e - N/2}{N/2} \tag{3}$$

where:

n_e is the number of experts who *agree* and *strongly agree*, N refers to the total number of experts.

Content Validity Ratio (CVR) is interpreted as, if more than half of the experts *agree* or *strongly agree* on the importance of an element/category then the CVR returns positive value ($CVR > 0$). On the other hand $CVR < 0$ if less than half do the same. The minimum value of CVR_{min} for the number of experts is calculated and given below, and only BMS elements that meet the minimum value can be included.

Table 5
Content Validity Ratio(CVR)

N	5	6	...	20	25	30	35	40
CVR_{min}	.99	.99	...	.42	.37	.33	.31	.29

Note. adopted from Lawshe (1975)

Reliability. The reliability of the survey is the degree of reproducibility of the results on a similar survey. GNU PSPP, a program for statistical analysis of sampled data was used(gnu.org, 2018), and Cronbach's alpha For the 8 behavioral markers items was .73 which is considered acceptable(Gliem & Gliem, n.d.).

3.2.3 Relative Importance of BMS. From the theory of training on the literature review, we have established that optimal design of a training program starts from the need and importance given to the training elements. Establishing a hierarchy of importance helps to prioritize training elements and evaluate the overall performance of training objective.

The Analytical Hierarchy Process (AHP) is a frequently used "multi-criteria decision-making approach"(Saaty, 1987) that quantifies expert opinion and tacit knowledge, and assigns an order of hierarchical, for ease of use in a decision making or evaluating process. Concerning Behavioral markers, this will be achieved by the following process.

Experts conduct a pairwise comparison of all elements on an importance scale where 1 denotes *equal importance* of both elements and 9 denotes *extreme importance* of one element/category over the other See table 22.

The mathematical representation is that the importance of element1(E1) compared to element2(E2) is the multiplicative inverse of that of E2 compared to E1. If importance of $E1/E2 = x$ then importance of $E2/E1 = 1/x$. For $E_1, E_2, ..E_n$ behavioral elements, then the number of comparisons that can be made are $n(n-1)/2$

resulting in a paired comparison matrix of n order.

$$E = \left[E_i/E_j\right] = \begin{bmatrix} E_1/E_1 & E_1/E_2 & \dots & E_1/E_n \\ E_2/E_1 & E_2/E_2 & \dots & E_2/E_n \\ \dots & \dots & \dots & \dots \\ E_n/E_1 & E_n/E_2 & \dots & E_n/E_n \end{bmatrix}$$

Where $A = [a_{ij}]$ denotes the preference intensity matrix of experts between behavioral elements, the relative significance of each elemenet/element/category is calculated by solving for normalized right-eigenvector.

$$A = \left[a_{ij}\right] = \begin{bmatrix} 1 & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ 1/a_{12} & 1 & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1/a_{1j} & 1/a_{2j} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \dots & a_{in} & \dots & 1 \end{bmatrix} \Rightarrow A \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \vdots \\ \upsilon_i \\ \vdots \\ \upsilon_n \end{bmatrix} = \lambda \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \vdots \\ \upsilon_i \\ \vdots \\ \upsilon_n \end{bmatrix}$$

which can be represented as

$$A\upsilon = \lambda\upsilon \tag{4}$$

Saaty (1987) defined *Consistency Index (CI)* for a matrix of order n as follows:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{5}$$

If a preference of an element is higher than a second element and the second element preference is higher than a third element, then the third element can not have preference higher than the first element. If the evaluation is perfectly consistent then $\lambda_{max} = n$ and $CI = 0$.

Pairwise comparisons are however influenced by human-factors, and therefore inconsistency is to be expected. *Consistency Ratio* for a given order of matrix is calculated using the *Random Consistency Index (RI)* shown at table 6, which is the

average CI of a 500 randomly generated matrices, which is expected to be highly inconsistent. $CR = CI/RI$ and may not exceed 0.10.

Saaty (1987) further notes inconsistency to be an important property that indicates a need for a readjustment of knowledge, which experience shows to be never consistent. Concerning BMS this can be used as an indication of the degree of disagreement between the experts and can be used as a starting point of discussion for *Shared Mental Model*(Jonker, van Riemsdijk, & Vermeulen, 2011) of NTS.

Table 6
Random Consistency Index(RI)

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Note. adopted from Saaty (1987)

For ease of use, we will employ an excel AHP calculation template implementation (revision 2018) based on the paper by (Goepel, 2013) see table 24 for credit.

3.3 Sampling strategy

A random sampling strategy of a group of respondents chosen for a survey, that represents the property of a larger population, employs a set of mathematical methods to avoid bias. (Creswell & Creswell, 2017) In qualitative research, difficulty to access the target population or the nature of the research may pose a problem, and random sampling may not be convenient. In such a case researchers often use a non-probability (convenience) sampling technique (Reddy & R., 2016), which will be used for this thesis.

Three car and passenger ferry shipping companies Human Resources(HR) departments were contacted by email, the purpose and method of the survey were explained, and assistance in identifying potential participants was requested. One company responded positively: for disclosure purpose, the researcher works for the said company as a marine engineer on board one of the car ferries. The companies HR department provided with 347 email addresses of Captains and chief-Officers from their

contact database, upon an *explicit request* for "Captains" and "chief officers", who in this case considered subject matter experts. Except for the officers that took part in the pilot survey and the AHP analysis, there was no direct contact between the researcher and the participants for the purpose of the survey.

Accessing a research participant indirectly through a third party requires that the eligibility of participants to be verified by a third party(Biernacki & Waldorf, 1981). Shipping companies keep up-to-date email-lists for group contact in each or several categories of employment, it can reasonably be assumed that the list is for the correct target subject matter experts, which sets the boundary of sampling. Moreover, a question about their position was included to make sure that the right target sample was obtained.

Potential problems include a perception of junk mail, unclear questionnaire and answering instructions, impersonality of electronic communication, representativeness and ensuring randomness within the boundary, due to low response rate can be taken as limitations on the sampling strategy (Evans & Mathur, 2005).

Inclusion criteria. The participating experts shall be bridge officers with at least four years of experience in a Norwegian merchant vessel.

3.4 Questionnaire Design

In research practice, it is recommended to use a previously used and tested methods and questionnaires, on the same subject, unless there is a compelling reason that a new method is required. The methodologies identified by literature review on similar NTS and BMS were resource and time consuming which is not available for this thesis. Therefore the author developed the questionnaires within the guidelines of sound questionnaire developing practice (Rattray & Jones, 2007; Sullivan & Artino, 2013).

Three-part questionnaires were designed with four objectives. The objectives of the first part questionnaires were to establish face validity of the elements identified in the literature review and to identify wording problems and clarity. This was done by *yes* and *no* answering and open-ended questions on comparing and grouping the

elements identified in the literature review.

The second part of the questionnaires was designed to establish content validity and reliability of the constructs and uses a five-point Likert-type scale, *strongly agree* to *strongly disagree* using the pilot questioner result as input. The second part of questionnaires also include five demographic questions; three to establish the respondent's expertise and experience, two questions designed to get an insight into the existing BRM training and utilization at the workplace.

The third part of the survey calculates the BMS relative importance by using the AHP tool developed by Goepel (2013).

3.5 Research Ethics

The EU *General Data Protection Regulation (GDPR)*(EU, 2018) places a strict requirement for online privacy and personal data protection. A detailed consent form was sent with an explanation and detailed objective of the research and a requirement to indicate an agreement was placed before submission. Steps are taken not to collect personal or demographic data that may identify a person more than the absolute necessity for the research, and those that are collected on the course of the research are kept confidential and were and will not be divulged to an outside party. The steps taken to ensure privacy was reviewed by the *Norwegian center for research Data(NSD)* and approved subsequently see document17.

The analysis and conclusions drawn in this research are drawn from the data collected during this study unless otherwise indicated.

4 Research results and Analysis

In this section, the results of the research process as described in the introduction and methodology section will be presented and the findings will be discussed. It consists of the results of the literature review, the findings of the validation of identified BMS, and the results of establishing the relative weight of BMS.

4.1 Results of the Literature review

In the literature review section the state of the art of learning, training and the underlining theoretical concepts were discussed, the role of evaluation was established, and methods were identified. The rationale behind the CRM training and evaluation from its inception in the aviation industry, evolution and adaptation into other safety-critical industries was discussed. BMS were identified as producing the best result for application in CRM training. The theoretical background and practical aspect of BMS were discussed and applicable BMS were extracted See table 17. For a more robust results the BMS extraction process were also identified. But due to resource limitation a decision was made to limit the effort on this thesis to test content validity of the identified BMS to Norwegian merchant shipping and calculate their relative importance.

4.1.1 BMS. Flin et al. (2003) has given the justification for each element on NOTECHS BMS and it will not be reproduced here. The limitation on this thesis required the reduction in a number of elements, by putting related elements together, effectively creating an intermediate category, to reduce complexity and minimize *cognitive burden*(Bowling, 2005). The authoritative definition for each construct thus is as given by Flin et al. (2003). This should also not affect the objective of addressing the research problem, that is establishing BMS as a feasible method to evaluate the effectiveness of BRM.

4.2 Pilot survey results

The questionnaire development process and re-categorization of the NOTECHS BMS into an intermediate category was done through two types of pilot surveys. The

Comparison of Non-technical Skills Categorization

	Warsash maritime academy	NOTECHS	NTS-maritime-bridge-no
Social Skills	**Co-operation**		
	Open communication	Supporting Others	Consideration to others: Supporting others
	Consideration for others	Considering Others	
	Team working	Team building maintaining	Team building: Team maintaining: Conflict solving:
		Conflict Solving	
	Leadership and managerial Skills		
	Assertiveness	Use of authority assertiveness	Use of authority: assertiveness
	Situational leadership	Providing and maintaining standards	
	Planning and coordinating	Planning and coordination	Planning: Coordination: Workload management
		Workload management	
Cognative Skills	**Situational awareness**		
	Situation assessment	Awareness of aircraft systems	Awareness of vessel: Internal systems: External Environment
	Risk assessment	Awareness of external environment	
		Awareness of time	Awareness of time, Risk assessment
	Decision making		
	Problem Diagnosis	Problem definition and diagnosis	Problem definition, diagnosis
	Option Generation	Option generation	Option generation: Selection: Outcome review
	Option selection	Risk assessment option selection	
		Outcome review	

Table 7
NTS Comparison table

Non-Technical Skills			
Social skills		Cognitive skills	
Categories	Elements	Categories	Elements
Cooperation	E1. Consideration to others, Supporting others	Situational Awareness	E5. Awareness of vessel Internal systems, Awareness of vessel External Environment
	E2. Team building, maintaining, Conflict solving		E6. Awareness of time Risk assessment
Leadership and managerial	E3. Use of authority, assertiveness	decision making	E7. Problem definition, diagnosis
	E4. Planning, coordination, Workload management		E8. Option generation,option selection, Outcome review

Table 8
Non-Technical Skills adopted from (P. M. Barnett et al., 2006; Devitt et al., 2010; Flin et al., 2003; O'Connor, 2011)

first part was a *yes* and *no* questionnaire and the second was an open-ended question to comment on the re-categorisation of the elements. The researchers role was to explain the elements concepts as described on NOTECHS, ask them how best they would re-categorize, and take note. The discussions have taken place intermittently in two work weeks on-board with in a six week period.

	Results of pilot survey		
Code	Category	Relevant %y	Ease %y
E1	Consideration to others, Supporting others	100%	100%
E2	Team building, maintaining, Conflict solving	100%	100%
E3	Use of authority, assertiveness	100%	100%
E4	Planning, coordination, Workload management	100%	100%
E5	Awareness of vessel Internal systems, Awareness of vessel External Environment	100%	100%
E6	Awareness of time, Risk assessment	100%	100%
E7	Problem definition, diagnosis	100%	100%
E8	Option generation,option selection, Outcome review	100%	100%

Table 9
Pilot survey results

Discussion. *Risk assessment* element overlaps many categorise, however there was agreement that the vessel being a scheduled car ferry its relatedness to *awareness of time* would be more appropriate on this study. Therefore it was agreed to categorize it

as such.

Open communication that is given it's own element by P. M. Barnett et al. (2006) is omitted in favor of the justification that *communication* to be a "general mediator" to all categories as pointed out by Flin et al. (2003). The results of the intermediate categorization and pilot survey are given on table 9. An agreement was reached on re-categorization of the elements as given on table 7.

4.3 Survey results

The survey was placed at an online survey site Zoho survey and the link was sent to the 347 email addresses in two groups. In the email a description of the survey, it's purpose and what is expected from the participants was written in Norwegian.

Four email address were reported not deliverable. The survey run for one week, between 11th and 18th of October 2018, with one reminder sent at the end of the fourth day.

At the conclusion of the survey 40 responses were registered an (11.7% response ratio). Since the assumption is, the respondents are experts with the same type of training and have similar experience in the same work environment, the response ratio is not considered an issue(Bowling, 2005).

Table 10
Participants statistics validation of BMS,

	Age	Years at sea	Years at position
mean	41.35	18.3	9.1
median	41	17	7
minimum	23	5	0
maximum	65	40	30
St.dev	11.5	11.4	8.4
n=40	captains=27	chief-officers=13,	
male=38,	female=2,		

The respondents were between 23 and 65 years old with a mean age of 41.35, while having a mean of 18.3 years experience at sea, out of which 9.1 years are served at the present position with minimum of 5 years of sea-time. 27 Captains and 13 chief-officers have taken part while only 2 out of 40 respondents were female see table 10.

Consensus and Content validity.												
Survey questions	SD	D	N	A	SA	mean	st.dev	consensus	A+SA	Total	%(A+SA)	CVR=(N(A+SA)-N/2)/N/2
E1: Consideration to others; Supporting others	0	0	5	22	13	4.21	0.67	0.80	35	40	87.50%	0.75
E2: Team building, maintaining, Conflict Solving	1	0	4	14	21	4.35	0.84	0.72	35	40	87.50%	0.75
E3:Use of authority, Assertiveness	2	2	4	23	9	3.84	1	0.71	32	40	80.00%	0.6
E4:Planning, coordination, Workload management	0	0	0	25	15	4.37	0.54	0.83	40	40	100.00%	1
E5:Awareness of vessel Internal systems; Awareness of vessel External Environment	0	0	4	16	20	4.37	0.69	0.78	36	40	90.00%	0.8
E6. Awareness of time; Risk assessment	0	0	4	21	15	4.26	0.66	0.80	36	40	90.00%	0.8
E7. Problem Diagnosis	0	0	6	23	11	4.09	0.65	0.82	34	40	85.00%	0.7
E8. Option generation; Option Selection	0	0	9	23	8	3.98	0.67	0.83	31	40	77.50%	0.55
E9: Bridge resource management training is essential.	1	2	6	21	10	3.92	0.92	0.74	31	40	77.50%	
E10: Bridge resource management training skills are implemented where I work.	2	2	10	22	4	3.6	0.93	0.71	26	40	65.00%	
Legend	SD=-2 Strongly disagree		D=-1 Disagree		N=0 Neutral		A=1 Agree		SA=2 Strongly agree			

Table 11
consensus and Content validity

Consensus. The percentage of respondents that *agree* and *strongly agree* on each category was between 77.7% and 100%. The consensus as given on equation (2) among the experts is between 0.71 and 0.83 where 1 is *complete agreement* and 0 is *complete disagreement. Use of authority, assertiveness* is the category that received the highest (4) *disagree* and *strongly disagree*, while receiving the second least number of (9) *strongly agree* response.

Content Validity Ratio (CVR). The CVR values of the survey between .6 and 1 are well above the minimum value of .29 for N=40 which indicates a strong consensus by the experts confirming all the categories as essential.

Reliability. The reliability test was carried out using Cronbach's alpha coefficient for the 8 behavioral markers was .73 where .7 is considered an acceptable

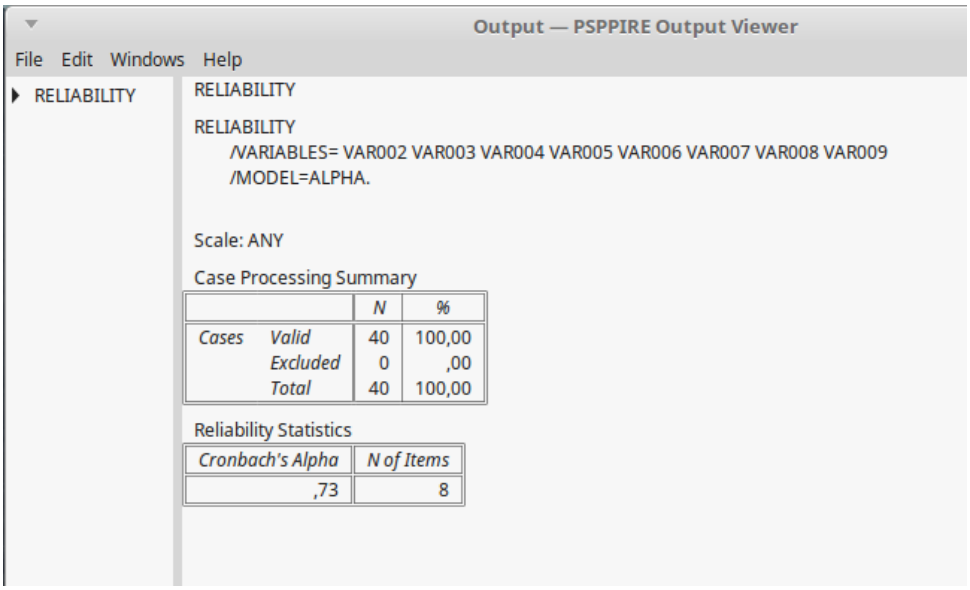


Table 12
Reliability

lower value. See table 12.

4.4 Analytical Hierarchy Process results

The initial AHP questionnaire submitted with the validation survey was abandoned as it has a high degree of inconsistency. Learning from the first questionnaire, a second questionnaire with better formulation was presented to 3 bridge officers who helped with the pilot survey. After three iterations a low enough consistency ratio of 6.6% was achieved.

AHP Analytic Hierarchy Process (EVM multiple inputs)

K. D. Goepel Version 15.09.2018 — Free web based AHP software on: http://bpmsg.com

Only input data in the light green fields and worksheets!

n=	8	Number of criteria (2 to 10)	Scale:	1	AHP 1-9
N=	3	Number of Participants (1 to 20)	a:	0.1	Consensus: 95.8%
p=	0	selected Participant (0=consol.)	2	7	Consolidated
Objective	Weighted hirarchy of beheioral markers.				
Author	Tegegne Tefera				
Date	30.03.2019	Thresh: 1E-08	Iterations: 5.0E+0(		EVM check: 3.8E-09

Table	Criterion	Comment	Weights	+/-
1	E1	consideration of others, Supporting Others:	3.4%	1.1%
2	E2	Team building, maintaining, conflict solving	7.5%	3.1%
3	E3	Use of authority, assertiveness	6.5%	3.0%
4	E4	Planning, coordination workload management	11.0%	4.7%
5	E5	Awareness of vessel internal systems, Awareness o	18.8%	11.3%
6	E6	Awareness of time, Risk management	19.9%	5.4%
7	E7	Problem definition and diagnosis	17.1%	8.6%
8	E8	Option generation, option selection, Outcome review	15.8%	4.8%
9		for 9&10 unprotect the input sheets and expand the	0.0%	0.0%
10		question section ("+" in row 66)	0.0%	0.0%

Result							
Eigenvalue			Lambda:	8.644		MRE:	42.4%
Consistency Ratio	0.37	GCI: 0.24	Psi: 8.9%	CR: 6.6%			

Table 13

Analytically hierarchy process for relative weight summery

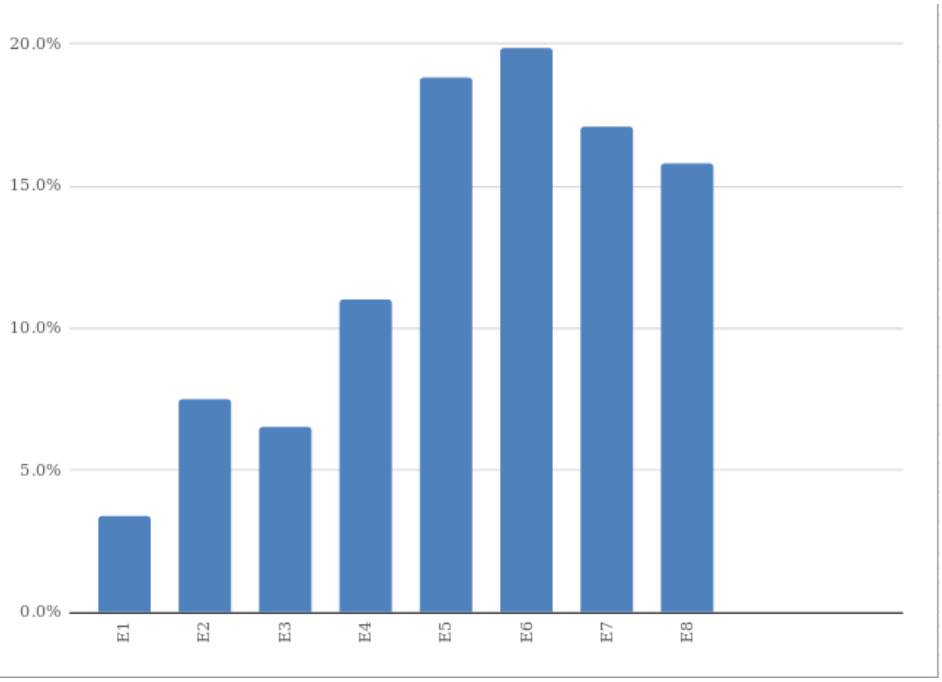


Table 14

Analytically hierarchy process for relative weight graphical

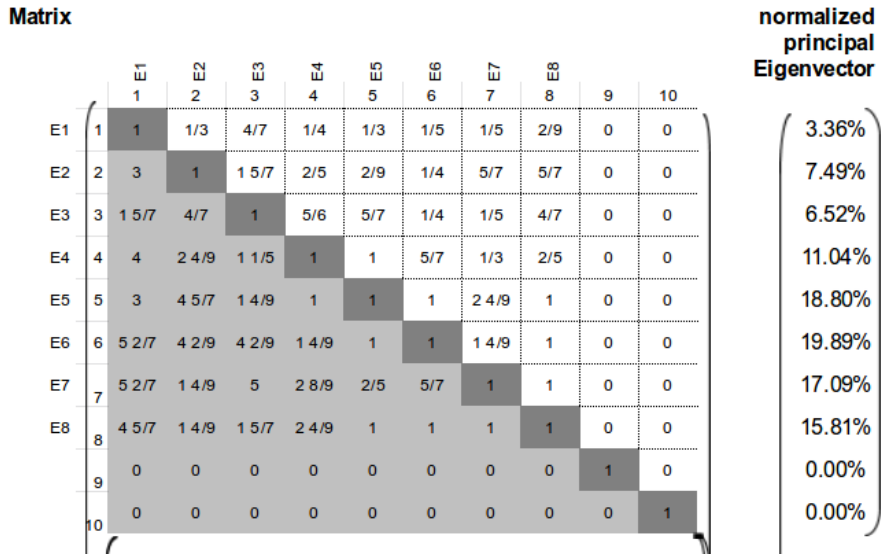


Table 15
Analytically hierarchy process for relative weight eigenvalue

		BMS category relative importance	
Rank	Code	Category	Relative importance
1	E6	Awareness of time, Risk assessment	19.9%
2	E5	Awareness of vessel Internal systems, Awareness of vessel External Environment	18.8%
3	E7	Problem definition, diagnosis	17.1%
4	E8	Option generation,option selection, Outcome review	15.8%
5	E4	Planning, coordination, Workload management	11.1%
6	E2	Team building, maintaining, Conflict solving	7.7%
7	E3	Use of authority, assertiveness	6.5%
8	E1	Consideration to others, Supporting others	3.4%

Table 16
relative weight

5 Discussions

5.1 Revisiting the research questions

The objective of this research was to find a BRM evaluation method that can better indicate whether *transfer* has occurred or not. The two first levels of Kirkpatrick's training evaluation methods namely the *Reaction* and *learning* can predict transfer or identify training needs, however, the characteristics of *safety-critical industries* necessitates a more certain measurement of behavior change as a result of training.

Behavioural Marker Systems are a taxonomy of key observable behaviors that can be measured to indicate a behavior change, thereby showing a training objectives are achieved. The development of BMS in safety-critical industries and its results were discussed, and the model for the Norwegian shipping sector was identified answering question number one; *What are the key Behavioral Marker Systems developed for safety-critical industries?*

5.2 The need and applicability of BMS

Two questions intended to get an insight into the state of BRM training implementation on-board were presented to the expert group.

- E9: Bridge resource management is essential.
- E10: Bridge resource management training skills are implemented where I work.

While 22.5% of respondents did not think the BRM training is essential 35% of respondents have the opinion that NTS training is not being implemented at their workplace. Which implies that there is a need to evaluate closely the BRM training and it's implementation in Norway to insure it's objectives are met.

The second research question *Do the behavioural markers apply to the Norwegian merchant ships bridge team?* was answered by demonstrating the content validity of the categories and the reliability of the survey. For a survey result to be valid for a number

of experts(n=40), the minimum content validity ratio is given .29. The calculated CVR value is between .6 and 1, which exceeded the minimum value.

The reliability test using Cronbach's alpha coefficient For the 8 behavioral markers items .73 where .7 is acceptable showed that the survey is reproducible under the same assumptions.

The consensus among experts on the relevance of each category to the Norwegian shipping sector was between 77.7% and 100%.

5.3 NTS priority.

The intent of Question number 3, *What is the relative importance of the behavioral markers?* was to align the derived BMS to the objectives of the Norwegian merchant shipping BRM training.

Hofstede, Hofstede, and Minkov (2010) defines culture as "the software of the mind". A more technical account of culture is given by R. Helmreich (2000) as "consisting of shared norms, values, and practices associated with a nation, organization, or profession". R. Helmreich (2000) further elaborates, that two of Höfsetedes cultural dimensions, *power distance*, and *individualism-collectivism* have particular relevance to CRM objectives.

The study revealed that *use of authority* and *Consideration to others, supporting others* categories scored the lowest relevance at 6.5% and 3.4% respectively. The results correlate with Hofstede's *power distance* and *individualism* scores for Norway which is placed relatively low (31) on *power distance*, and high(69) on *individualism*.

The categories under *Situational Awareness* received the highest scores 19.9% and 18.8%. Situational Awareness (SA).

The survey reveals that *Leadership and Managerial* and *Decision-making* skills are less prioritized than *Situational Awareness* skills, while *Cognitive skills* are more valued than *Social skills*.

While preparing for this thesis, the author discussed with the head of the BRM training at the University of South-East Norway on the appropriateness of the BRM

training material for the Norwegian context. The response was that the material is too generic and “was not prepared having Norwegian seafarers in mind”. A generic one-fit-for-all type of CRM training design discounts the cultural difference of the trainees and BMS attempts to address this missing element(R. Helmreich, 2000) by identifying the specific training needs of the trainee.

6 Conclusion

The Norwegian frigate KNM Helge Ingstad and a commercial oil tanker Sola TS collided near the Sture terminal in Hjeltefjorden, Norway. The frigate sustained extensive damage and abandoned, sinking close to shore. The Accident Investigation Board Norway (AIBN) and the Defense Accident Investigation Board Norway (DAIBN) issued a preliminary marine accident report, On 29th November 2018, and stated that they did not find any technical systems that "did not function" up until the accident moment. They also suggested a further line of investigation to be "human factors, collaboration on the bridge, training and procedures, traffic control, language and communication"(AIBN & DAIBN, 2018) among others, which is a subject this research is trying to address. Though it is early to make any conclusion, the accident puts a spotlight on the resource management training and evaluation regime of both the military and civilian maritime industry in Norway.

Figure 5. KNM Helge Ingstad

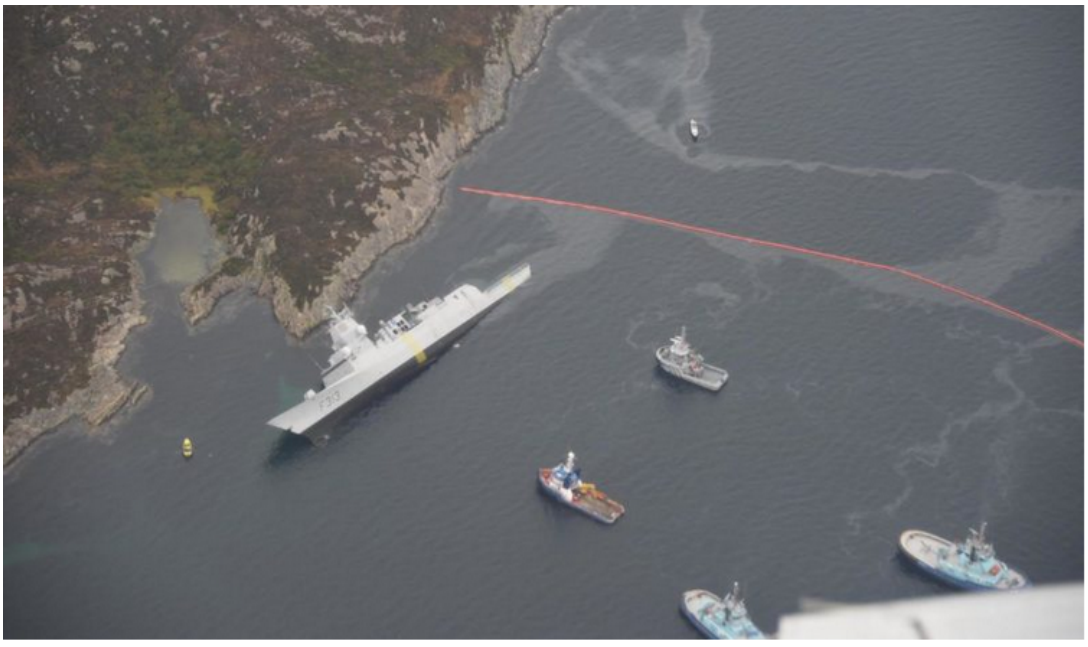


photo. Norwegian Coastal Authority

The Norwegian maritime authority issued guidelines (Sjøfartsdirektoratet, 2017) for the circular on BRM and ERM requirement(*Krav Til BRM- Og ERM-Kompetanse for Dekks- Og Maskinoffiserer i Henhold Til STCW-78-Konvensjonen, Med Endringer -*

Sjøfartsdirektoratet, 2014), and the possible avenues to obtain the approved competency, and the requirements for assessors are given in these documents. It can be seen from this documents that there is no special training requirement for assessors as it is for BMS raters that are employed in CRM training evaluations. The adequateness of this practice should be studied and compared with the practice in other safety-critical industries.

The author as part of to fulfill the requirement for a marine engineer certificate attended an Engine Resource Management (ERM) training that is conducted in tandem with BRM at USN facility in Horton. Moreover, the author attended a one-day Maritime Resource Management (MRM) facilitator training arranged by ALL Academy International AB, in Gothenburg Sweden, as a preparation for this research. In both cases, the author noticed that the training evaluation is limited to Kirkpatrick's reactions levels, which according to research is a poor predictor of transfer.

An additional issue worth mentioning as M. Barnett et al. (2003) points out is that the IMO assessment model does not differentiate between emergency and crisis management. An emergency is an undesirable incident that has a pre-defined procedure for remedy, while a crisis has no such procedure. These differences affect the training and assessment requirements.

6.1 Research Contribution

BMS development and validation is an ongoing research effort for safety-critical industries and their application in aviation, operating theaters, nuclear and rail industries hitherto had a positive effect on safety(Klampfer et al., 2001).

This research has

- demonstrated that there is a need for a better evaluation tool for BRM training for Norwegian merchant shipping,
- drew attention to BMS to be used as a training and competency evaluation tool,
- showed that the existing work on BMS can be utilized as a starting point to develop BMS for the Norwegian merchant shipping sector,

- identified the priorities of categories and elements of BMS as identified in this research and applied to the cultural context sampled bridge team.

6.2 Research limitations

The development of behavioural markers is expensive and requires the collaboration of multidisciplinary domain knowledge experts and industry (Klampfer et al., 2001). This research lacked the resources, data, and subject matter experts with the experience and knowledge, therefore unable to employ the recommended methodology that the study requires.

The sample was taken only from one type of shipping sector, therefore, can not be generalised to all shipping sector in Norway.

6.3 Suggestion for future research

BMS are continually evolving, responding to continually changing technical solutions and corresponding task requirements. Future research on the topic should follow the recommendations made by Flin and O'Connor (2017); Klampfer et al. (2001) by analyzing data of "accident investigation, confidential incident reporting systems, incident analysis, simulator studies, task analysis, interviews, surveys, focus groups, ethnographies" and the performances that "contributed to successful and unsuccessful outcomes".

This study can be extended by

- developing an appropriate scoring method for the behavioral markers,
- developing appropriate simulator scenario to observe the task performance,
- testing the validity of the results in relation to performance of trainees,
- test internal consistency and inter-rater reliability.

7 References

AIBN, & DAIBN. (2018). *Preliminary report collision between frigate KNM helge ingstad and the oil tanker sola TS 8 november 2018.pdf.* Accident Investigation Board Norway. Retrieved from `https://www.aibn.no/Marine/Investigations/18-968` (00000)

Aldoobie, N. (2015, December). ADDIE Model. *American International Journal of Contemporary Research*, *Vol. 5*(No. 6).

ALL Academy. (2019). *Maritime Resource Management (MRM) - formerly Bridge Resource Management (BRM).* (00000)

Allen, W. C. (2006, November). Overview and Evolution of the ADDIE Training System. *Advances in Developing Human Resources*, *8*(4), 430-441. doi: 10.1177/1523422306292942

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997, June). A META-ANALYSIS OF THE RELATIONS AMONG TRAINING CRITERIA. *Personnel Psychology*, *50*(2), 341-358. (01258) doi: 10/dn84s7

Alvarez, K., Salas, E., & Garofano, C. M. (2004, December). An Integrated Model of Training Evaluation and Effectiveness. *Human Resource Development Review*, *3*(4), 385-416. doi: 10.1177/1534484304270820

Amalberti, R. (2001, March). The paradoxes of almost totally safe transportation systems. *Safety Science*, *37*(2), 109-126. (00577) doi: 10/dgd3cw

Arora, S., Miskovic, D., Hull, L., Moorthy, K., Aggarwal, R., Johannsson, H., ... Sevdalis, N. (2011). Self vs expert assessment of technical and non-technical skills in high fidelity simulation. *The American Journal of Surgery*, *202*(4), 500 - 506. doi: https://doi.org/10.1016/j.amjsurg.2011.01.024

Baldwin, T. T., & Ford, J. K. (1988). Transfer of Training: A Review and Directions for Future Research. *Personnel Psychology*, *41*(1), 63-105. (03777)

Barnett, M., Gatfield, D., & Pekcan, C. (2003). A Research Agenda in Maritime Crew Resource Management. , 23. (00023)

Barnett, P. M., Gatfield, D., & Pekcan, C. (2006). Non-technical skills: The vital

ingredient in world maritime technology? , 11. (00018)

Biernacki, P., & Waldorf, D. (1981, November). Snowball Sampling: Problems and Techniques of Chain Referral Sampling. *Sociological Methods & Research*, *10*(2), 141-163. (03998) doi: 10/gd3z6z

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010, July). Transfer of Training: A Meta-Analytic Review. *Journal of Management*, *36*(4), 1065-1105. (00832) doi: 10/ftwgxk

Bowling, A. (2005, September). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*(3), 281-291. (01274) doi: 10/fjsrs3

Bužančić Primorac, B., & Parunov, J. (2016). Review of statistical data on ship accidents. In C. Guedes Soares & I. A. Santos (Eds.), *Maritime Technology and Engineering* (p. 809-814). London: Taylor. (00005)

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J.-P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis and Prevention*, *59*, 26–37. (00131) doi: 10.1016/j.aap.2013.05.006

Crayton, L., Hackworth, C., Roberts, C. A., & King, S. J. (2001). Line Operations Safety Assessments (LOSA) in Maintenance and Ramp Environments. , 48. (00000)

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Devitt, K., Holford, S., Pantaleev, B., & Sharma, D. (2012, October). The validation of non-technical behavioural markers for Merchant Navy officers. In *5th International Conference on Maritime Human Resource Solutions*. St John's, Newfoundland, Canada. (00000)

Devitt, K., Hons, J., & Holford, S. (2010). The development of resource management and leadership behavioural markers for the Merchant Navy. , 24. (00009)

Dietrich, R., & Chhildress, T. M. (2017). *Group Interaction in High Risk*

Environments. Routledge. (00036)

Emsa. (2017). *Annual overview of marine casualties and incidents 2017_final.pdf*. European Maritime Safety Agency. (00000)

EU. (2018, July). *Data protection*. https://europa.eu/youreurope/citizens/consumers/internet-telecoms/data-protection-online-privacy/index_en.htm. (00271)

Evans, J. R., & Mathur, A. (2005, April). The value of online surveys. *Internet Research*, *15*(2), 195-219. (01619) doi: 10/d8tcnt

Flin, R., Martin, L., Goeters, K.-M., Hörmann, H.-J., & Amalberti, R. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. , 23. (00252)

Flin, R., & O'Connor, P. (2017). *Safety at the Sharp End: A Guide to Non-Technical Skills*. CRC Press. (00971)

Flin, R., O'Connor, P., & Mearns, K. (2002, June). Crew resource management: Improving team work in high reliability industries. *Team Performance Management: An International Journal*, *8*(3/4), 68-78. (00194) doi: 10/fwqsm6

Flynn, F. M., Sandaker, K., & Ballangrud, R. (2017, January). Aiming for excellence – A simulation-based study on adapting and testing an instrument for developing non-technical skills in Norwegian student nurse anaesthetists. *Nurse Education in Practice*, *22*, 37-46. (00000) doi: 10/f9vsbr

F. P. Da Conceição, V., Basso, J. a., Lopes, F. C., & Dahlman, J. (2017). Development of a Behavioural Marker System for Rating Cadet ' s Non- Technical Skills. *Transnav, International Journal On Marine Navigation And Safety Of Sea Transportation*, *11*(2). (00000) doi: 10/gcjncq

Garavan, T. N. (1997). Training, development, education and learning: Different or the same? *Journal of European Industrial Training*, *21*(2), 39–50. doi: 10.1108/03090599710161711

Gliem, J. A., & Gliem, R. R. (n.d.). Calculating, Interpreting, and Reporting

Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. , 7. (03208)

gnu.org. (2018). *Pspp.* https://www.gnu.org/software/pspp/. (00003)

Goepel, K. D. (2013). IMPLEMENTING THE ANALYTIC HIERARCHY PROCESS AS A STANDARD METHOD FOR MULTI-CRITERIA DECISION MAKING IN CORPORATE ENTERPRISES – A NEW AHP EXCEL TEMPLATE WITH MULTIPLE INPUTS. , 10. (00107)

Guba, E. C. (1990). *The alternative paradigm dialog . In E. G. Guba (Ed.), The paradigm dialog (pp. 17 – 30).* SAGE Publications.

Hansen, H. L. (2002, February). Occupational accidents aboard merchant ships. *Occupational and Environmental Medicine*, *59*(2), 85-91. (00103) doi: 10/c4r2pb

Helmreich, R. (2000). Culture, Threat, and Error:. In *Assessing system safety. In Safety in Aviation: The Management Commitment* (p. 12). London. (00072)

Helmreich, R. L. (1997, May). Managing Human Error in Aviation. *Scientific American*, *276*(5), 62-67. doi: 10.1038/scientificamerican0597-62

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (1999). Models of Threat, Error, and CRM in Flight Operations. In *Proceedings of the Tenth International Symposium on Aviation Psychology* (p. 677-682). Columbus, OH: The Ohio State University. (00183)

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999, January). The Evolution of Crew Resource Management Training in Commercial Aviation. *The International Journal of Aviation Psychology*, *9*(1), 19-32. doi: 10.1207/s15327108ijap0901_2

Hofstede, G. (2011, December). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1). doi: 10.9707/2307-0919.1014

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind ; intercultural cooperation and its importance for survival* (Rev. and expanded 3. ed ed.). New York: McGraw-Hill. (00417 OCLC: 699216499)

Illeris, K. (2003, July). Towards a contemporary and comprehensive theory of learning.

International Journal of Lifelong Education, *22*(4), 396-406. doi: 10.1080/02601370304837

IMO. (1995a). *International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW).* http://www.imo.org/en/About/Conventions/ListOfConventions/Pages/International-Convention-on-Standards-of-Training,-Certification-and-Watchkeeping-for-Seafarers-(STCW).aspx.

IMO. (1995b). *STCW Home - Standards of Training, Certification and Watchkeeping for Seafarers.* http://www.navit.fo/stcw%20kodan/stcw-code-atoc.htm. (00000)

IMO. (2014). *Leadership and team work Module Cource 1:39.* (00000)

JARTEL_Consortium. (2002, October). *Final Report JAR TEL Consolidation of Results.* https://trimis.ec.europa.eu/sites/default/files/project/documents/jartelrep.pdf. (00000)

Jarvis, P. (2012a). *Adult Education and Lifelong Learning: Theory and Practice.* Routledge.

Jarvis, P. (2012b). *Towards a Comprehensive Theory of Human Learning.* Routledge.

Jonker, C. M., van Riemsdijk, M. B., & Vermeulen, B. (2011). Shared Mental Models. In M. De Vos, N. Fornara, J. V. Pitt, & G. Vouros (Eds.), *Coordination, Organizations, Institutions, and Norms in Agent Systems VI* (p. 132-151). Springer Berlin Heidelberg. (00091)

Kirkpatrick, D. (2007). *The Four Levels of Evaluation.* American Society for Training and Development.

Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *Kirkpatrick's Four Levels of Training Evaluation.* Association for Talent Development. (00218)

Klampfer, B., Flin, R., & Helmreich, R. (2001). *Enhancing performance in high risk environments: Recommendations for the use of behavioural markers. Report from the behavioural markers workshop.*

Klinect, J., Murray, P., Merritt, A., & Helmreich, R. (2003). LINE OPERATION SAFETY AUDITS (LOSA): DEFINITION AND OPERATING CHARACTERISTICS. , 7. (00000)

Knowles, M. S. (1978, January). Andragogy: Adult Learning Theory in Perspective. *Community College Review*, *5*(3), 9-20. doi: 10.1177/009155217800500302

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, *78*(2), 311.

Krav til BRM- og ERM-kompetanse for dekks- og maskinoffiserer i henhold til STCW-78-konvensjonen, med endringer - Sjøfartsdirektoratet. (2014). https://www.sdir.no/sjofart/regelverk/rundskriv/krav-til-brm-og-erm-kompetanse-for-dekks-og-maskinoffiserer-i-henhold-til-stcw-78-konvensjonen-med-endringer/. (00000)

Laker, D. R., & Powell, J. L. (2011, March). The differences between hard and soft skills and their relative impact on training transfer. *Human Resource Development Quarterly*, *22*(1), 111-122. doi: 10.1002/hrdq.20063

Lawshe, C. H. (1975, December). A QUANTITATIVE APPROACH TO CONTENT VALIDITY. *Personnel Psychology*, *28*(4), 563-575. (03473) doi: 10/ccdqkv

Lowe, A., Hayward, B., & Branford, K. (2016). Leadership in safety critical industries: Project Report 1. , 75. (00000)

Masadeh, M. (2003). TRAINING, EDUCATION, DEVELOPMENT AND LEARNING: WHAT IS THE DIFFERENCE? , *8*, 7.

Merrill, M. D. (2002, September). First principles of instruction. *Educational Technology Research and Development*, *50*(3), 43-59. (02112) doi: 10/ckdfbf

Nielsen, D., & Roberts, S. (1999, January). Fatalities among the world's merchant seafarers (1990–1994). *Marine Policy*, *23*(1), 71-80. (00047) doi: 10.1016/S0308-597X(98)00017-7

O'Connor, P. (2002). Techniques used to evaluate Crew Resource Management

training: A literature review.

O'Connor, P. (2011, October). Assessing the Effectiveness of Bridge Resource Management Training. *The International Journal of Aviation Psychology*, *21*(4), 357-374. doi: 10.1080/10508414.2011.606755

O'Connor, P., Hörmann, H.-J., Flin, R., Lodge, M., Goeters, K.-M., & JARTEL Group, T. (2002, July). Developing a Method for Evaluating Crew Resource Management Skills: A European Perspective. *The International Journal of Aviation Psychology*, *12*(3), 263-285. (00100) doi: 10/br6p3f

O'Connor, P., & Max Long, W. (2011, December). The development of a prototype behavioral marker system for US Navy officers of the deck. *Safety Science*, *49*(10), 1381-1387. (00010) doi: 10/fqp7sh

Øvergård, K. I., Sorensen, L. J., Hontvedt, M., & Nazir, S. (2017). Maritime Bridge Crew Training. In M. S. Young & M. G. Lennè (Eds.), *Simulator for Transportation Human Factors: Research and Practice.* CRC Press LLC.

Patey, R., Flin, R., Fletcher, G., Maran, N., Glavin, R., & US Department of Defense; Defense Technical Information Center. (2005). *Developing a Taxonomy of Anesthetists' Nontechnical Skills (ANTS): (448192006-001).* American Psychological Association. (00000) doi: 10.1037/e448192006-001

Rattray, J., & Jones, M. C. (2007, February). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, *16*(2), 234-243. (00827) doi: 10/dxsscr

Reason, J. (2000, March). Human error: Models and management. *BMJ : British Medical Journal*, *320*(7237), 768-770. (05220)

Reddy, L. S., & R., K. (2016, July). Justifying The Judgmental Sampling Matrix Organization in Outsourcing Industry. *GBAMS- Vidushi*, *8*(02). (00000) doi: 10/gffjdt

Ritzmann, S., Hagemann, V., & Kluge, A. (2014). The Training Evaluation Inventory (TEI)–Evaluation of Training Design and Measurement of Training Outcomes for Predicting Training Success. *Vocations and Learning*, *7*(1), 41–73. doi:

10.1007/s12186-013-9106-4

Saaty, R. W. (1987, January). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, *9*(3), 161-176. (01346) doi: 10/csv75m

Saeed, F., Wall, A., Roberts, C., Riahi, R., & Bury, A. (2017, January). A proposed quantitative methodology for the evaluation of the effectiveness of Human Element, Leadership and Management (HELM) training in the UK. *WMU Journal of Maritime Affairs*, *16*(1), 115-138. (00000) doi: 10/gc957x

Salas, E., Rhodenizer, L., & Bowers, C. A. (2000, September). The Design and Delivery of Crew Resource Management Training: Exploiting Available Resources. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(3), 490-511. doi: 10.1518/001872000779698196

Salkind, N. (2010). *Encyclopedia of Research Design, Content Validity.* 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. (00002) doi: 10.4135/9781412961288

Schrader, P. G., & Lawless, K. A. (2004, September). The knowledge, attitudes, & behaviors approach how to evaluate performance and learning in complex environments. *Performance Improvement*, *43*(9), 8-15. (00067) doi: 10/c3dfbz

Sjøfartsdirektoratet. (2017). *Veiledning for bedømmelse av BRM- og ERM kompetanse -.* https://www.sdir.no/veiledninger/veiledning-for-bedommelse-av-brm–og-erm-kompetanse/. (00000)

Sullivan, G. M., & Artino, A. R. (2013, December). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, *5*(4), 541-542. doi: 10/gcvkds

Tastle, W. J., & Wierman, M. J. (2007, August). Consensus and dissention: A measure of ordinal dispersion. *International Journal of Approximate Reasoning*, *45*(3), 531-545. (00109) doi: 10/cgr79j

Thomas, M. J. W. (2017). *Training and Assessing Non-Technical Skills: A Practical Guide.* CRC Press. (00003)

Thorn, D. W., & Deitz, J. C. (1989, November). Examining Content Validity through the Use of Content Experts. *The Occupational Therapy Journal of Research*, *9*(6), 334-346. (00043) doi: 10/gfx6t7

Twitchell, S., Holton, E. F., & Trott, J. W. (2000, September). Technical Training Evaluation Practices in the United States. *Performance Improvement Quarterly*, *13*(3), 84-109. (00093) doi: 10/bcft2j

United Nations Conference on Trade and Development. (2017). *Review of maritime transport 2017*. (00000 OCLC: 1022725798)

Warr, P., Allan, C., & Birdi, K. (1999, September). Predicting three levels of training outcome. *Journal of Occupational and Organizational Psychology*, *72*(3), 351-375. doi: 10.1348/096317999166725

Wu, Y., Miwa, T., Shimamoto, K., & Uchida, M. (2015, October). Development of quantitative team performance evaluation method for ERM. *WMU Journal of Maritime Affairs*, *14*(2), 333-347. (00003) doi: 10/gc8zkk

Yilmaz, K. (2013). Comparison of Quantitative and Qualitative Research Traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, *48*(2), 311–325. (00650) doi: 10/gdj8qr

:

Non-Technical Skills			
Social skills		Cognitive skills	
Categories	Elements	Categories	Elements
Cooperation	E1. Consideration to others	Situational Awareness	E8. Awareness of vessel Internal systems
	E2. Team building and maintaining		E9. Awareness of vessel External Environment
	E3. Supporting others		E10. Awareness of time
	E4. Conflict solving		E11. Risk assessment
Leadership and managerial	E5. Use of authority,	decision making	E12. Problem Diagnosis
	E6. Planning and coordination		E13. Option generation
	E7. Workload management		E14. Option selection
			E15. Outcome Review

Table 17
Non-Technical Skills adopted from (Devitt et al., 2010; Flin et al., 2003; O'Connor, 2011)

Non-Technical Skills and Exemplar Behaviours-Social skills- Leadership and managerial skills Defined as "effective leadership and managerial skills achieve joint task completion within a motivated, fully functioning team through coordination and Persuasion" (Flin, Martin, Goeters, Hörmann, & Amalberti, 2003)		
	Behavioural Markers	
Elements	Good Behaiours	Poor Behaviours
Use of authority, Assertiveness	Initiates to involve team	Hinders involvement of team
	Takes command when needed	Passive, indecisive
	Assert own position	Unrecognizable position
	Reflects on suggestions	Ignores suggestion
	Appreciation	No appreciation
	Offers motivation	Offers no motivation
	Adequate coaching	Too little or too much coaching
Planning and coordination	Involve team in planning	Does not involve team
	Plans clearly stated and confirmed:	Plans not stated and confirmed:
	Goals, boundaries, task completion clearly stated	Goals, boundaries, task completion not clearly stated
	Accept change of plan if necessary	Too rigid to follow plan
	Consults team on change of plan	Change plan without consultation
Workload management	Appropriate task delegation	No or inappropriate task delegation or distribution:
	Appropriate task Prioritisation	Trivial tasks prioritised
	Allot adequate time to task:	Workload increase due insufficient time:
	Notes signs of stress, fatigue:	Ignores signs of stress, fatigue:

Table 18
Leadership and managerial skills NTS adopted from (M. Barnett et al., 2003; Devitt et al., 2010; O'Connor, 2011)

Non-Technical Skills and Exemplar Behaviours-Cognitive skills-decision making		
	Behavioural Markers	
Elements	Good Behaiours	Poor Behaviours
Problem Diagnosis	Gathers information, identify problem	Fails to diagnose, Problem not stated
	Causal factor appraisal with team	No discussion
Option generation	Elucidate alternatives	Fail to seek for alternatives
	Involves crew in alternative generation	Does not involve crew
Option selection	Discuss estimated risks of alternatives with team	Inadequate discussion of risk of alternatives with crew
	Crew limitations considered and discussed	No consideration of crew limitation
	Confirms and states selected option.	Crew is not informed of the rational of the alternative.
	Assess outcome against plan	No assessment of outcome against plan

Table 19
Decision making NTS adopted from (M. Barnett et al., 2003; Devitt et al., 2010; O'Connor, 2011)

Non-Technical Skills and Exemplar Behaviours-Cognitive skills- Situational Awareness Defined 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" (Flin, Martin, Goeters, Hörmann, Amalberti, 2003)		
	Behavioural Markers	
Elements	Good Behaiours	Poor Behaviours
Awareness of vessel Internal systems	Monitors and acknowledges system entries and changes	Does not follow updates
Awareness of vessel External Environment	Monitors external environment change	Not cognizant of external environment change
	Acknowledge environmental change	Not signal awareness of change
	Elucidate the information with the team	Does not share or seek to be informed
	Maintain and update outside resources	Poor maintenance of outside resources
Awareness of time	Discuss constraints and contingency with team	Does not discuss past, present and possible future events with the team.
	Cognizant of time for emergency procedures and crisis situation.	Do not show recognition of constraints
Risk assessment	Identify possible future problems	Shows surprise at out come of past events.

Table 20
Situational Awareness NTS adopted from (M. Barnett et al., 2003; Devitt et al., 2010; O'Connor, 2011)

Non-Technical Skills and Exemplar Behaviours-Social skills -Cooperation Defined as "The ability to work effectively in team" (Flin, Martin, Goeters, Hörmann, Amalberti, n.d.).		
	Behavioural Markers	
Elements	Good Behaiours	Poor Behaviours
Consideration others	Takes suggestions of others	Ignores suggestions
	Accounts for condition of others	Does noto account for condition of others
	Gives feedback	No feedback
Team building and maintaining	Establish Open communication,	Block open communication
	encourage inputs and feedback	keep distance from/between crew
	non-competitive	Competitive
Supporting others	Helps others in need	Hesitant to help
	Offers assistance	Does not offer to assist
Conflict solving	Keeps calm	Overreaction
	Suggest solution	Offer no compromise
	Focus on right solutions	Offers no solution
	Not focus on blame	Focus on blame

Table 21
Cooperation NTS adopted from (M. Barnett et al., 2003; Devitt et al., 2010; O'Connor, 2011)

The Fundamental scale

Intensity of imprtance on an absoulute scale	Definition	Explanation
1	Equal importance	The 1st skill is is equally important to the 2nd skill
3	Moderately important	Experience and judgment moderately favor the 1st skill over the 2nd
5	Strongly important	Experience and judgment strongly favor the 1st skill over the 2nd.
7	Very strongly important	The 1st skill is very strongly favored and it's dominance demonstrated in practice
9	Extremely important	The evidence favoring the 1st skill over the 2nd is of the highest possible order of affirmation.
2,4,6,8	Intermediate values	When compromise is needed
Reciprocals	If E1 is assigned a value when compared to E2 then E2 has a reciproacl value when compared to E1.	When compromise needed

Table 22
AHP fundamental scale adopted from (Saaty, 1987)

Table 23
Pilot Survay

Pilot survey

Comparison of Non-technical Skills Categorization		
Warsash maritime academy	NOTECHS	NTS-maritime-bridge-no
Co-operation	Supporting Others	
Social Skills	Considering Others	
Team working	Team building maintaining	
	Conflict Solving	
Leadership and managerial Skills	Use of authority assertiveness	
Assertiveness	Providing and maintaining standards	
Situational leadership	Planning and coordination	
Planning and coordinating	Workload management	
Situational awareness	Awareness of aircraft systems	
Cognitive Skills	Awareness of external environment	
Situation assessment	Awareness of time	
Risk assessment		
Decision making	Problem definition and diagnosis	
Problem Diagnosis	Option generation	
Option Generation	Risk assessment option selection	
Option selection	Outcome review	

Open ended question.
How would you re-categorize the elements?

	Element	Relevant	Easy
E1	Consideration to others, Supporting others		
E2	Team building, maintaining, Conflict solving		
E3	Use of authority, assertiveness		
E4	Planning, coordination, Workload management		
E5	Awareness of vessel Internal systems, Awareness of vessel External Environ men		
E6	Awareness of time, Risk assessment		
E7	Problem definition, diagnosis		
E8	Option generation,option selection, Outcome review		

Relevant -> y=yes; n=no
Easy -> y=yes; n=no

Document 1: NTS Requirements survey

Norwegian Maritime Bridge Officers Non-Technical Skills(NTS) requirements

Non-Technical Skills (NTS) requirements

University of South-Eastern Norway

Introduction

Eighty percent of all global trade and 70% of its value is transported by ships. Though it is considered one of the safest and economical transportation medium accidents and incidents happen and human factors are identified as fully or partially responsible for 80 to 85% of this incidents and accidents.

Companies send their officers to IMO mandated BRM/ERM training and spend millions to mitigate this accidents and incidents. Yet there is no reliable method to measure whether this training achieved its objective. Without such measurement, it is not possible to improve the training or justify the investment made.

The purpose of this survey is to obtain an expert opinion on Non-Technical Skills requirements for Norwegian Bridge officers as envisioned by STCW Convention Manila amendments of 2010 detailed in Table A-II/I Bridge Resource Management.

The data will be used to develop a reliable evaluation method for the effectiveness of a Bridge Resource Management training by observing on the job or simulated scenario of team performance to improve training design. It is not an individual performance evaluation tool.

It is voluntary to participate in the project, and you can at any time choose to withdraw your consent without stating any reason. If you decide to withdraw, all your personal data will be made anonymous.

All personal information is confidential and subject to the Norwegian data protection laws, regulations and research ethics.

Contact information

Researcher Tegegne Tefera

email: 012557@usn.no

phone: +47 95767154

Url: https://www.usn.no/

Document 2: NTS Requirements continued

Survey Guide

The survey is divided in two sections.

The first section is to establish the non-technical skills requirements for a Norwegian bridge officer to safely navigate a ship and dispense her/his duty as an officer.

The second section is to determine the relative importance of each non-technical skill identified, for a purpose of performance evaluation. This is done by comparing each skill with another skill.

Thank you for your participation.

Part 1 Instructions

Please choose the degree you agree or disagree with the statement.

When your answers to this questioner are registered and submitted you will be taken to the 2nd survey.

Q.1: How would you feel about the following statement.

Bridge resource management training is essential.*

- [-2] Strongly disagree
- [-1] Disagree
- [0] Neutral
- [1] Agree
- [2] Strongly agree

Document 3: NTS Requirements continued

Q.2: How would you feel about the following statement.

Bridge resource management training skills are implemented where I work.*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Q.3: The following cooperation skills are required:

Consideration of others; Supporting others;*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Document 4: NTS Requirements continued

Q.4: The following cooperation skills are required:

Team building, maintaining, Conflict Solving:*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Q.5: The following Leadership and managerial skills are required:

Use of authority, Assertiveness*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Document 5: NTS Requirements continued

Q.6: The following Leadership and managerial skills are required:
Planning, coordination, Workload management*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Q.7: The following Situational Awareness skills are required:

Awareness of vessel Internal systems;
Awareness of vessel External Environment*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Document 6: NTS Requirements continued

Q.8: The following Situational Awareness skills are required:

Awareness of time; Risk assessment*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Q.9: The following Decision making skill is required:

Problem Diagnosis*

-2 Strongly disagree

-1 Disagree

0 Neutral

1 Agree

2 Strongly agree

Document 7: NTS Requirements continued

Q.10: The following Decision making skills are required:

Option generation; Option Selection*

- [ ] -2 Strongly disagree
- [ ] -1 Disagree
- [ ] 0 Neutral
- [ ] 1 Agree
- [ ] 2 Strongly agree

Your position*

- [ ] Captain
- [ ] Chef officer
- [ ] Other (Please specify)

Number of years at sea*

Number of years at present position*

Year of Birth*

Document 8: NTS Requirements continued

Sex*

☐ Male ☐ Female

Submit

Please submit and continue to the next part of the survey.

By submitting this form you are indicating that you have read the description of the study, are over the age of 18, and that you agree to the terms as described.

Document 9: NTS Weights

Non-Technical skills relative weight.

1 = Equally important: Options are equally important

3 = Moderately important. Experience and judgment moderately favors option X

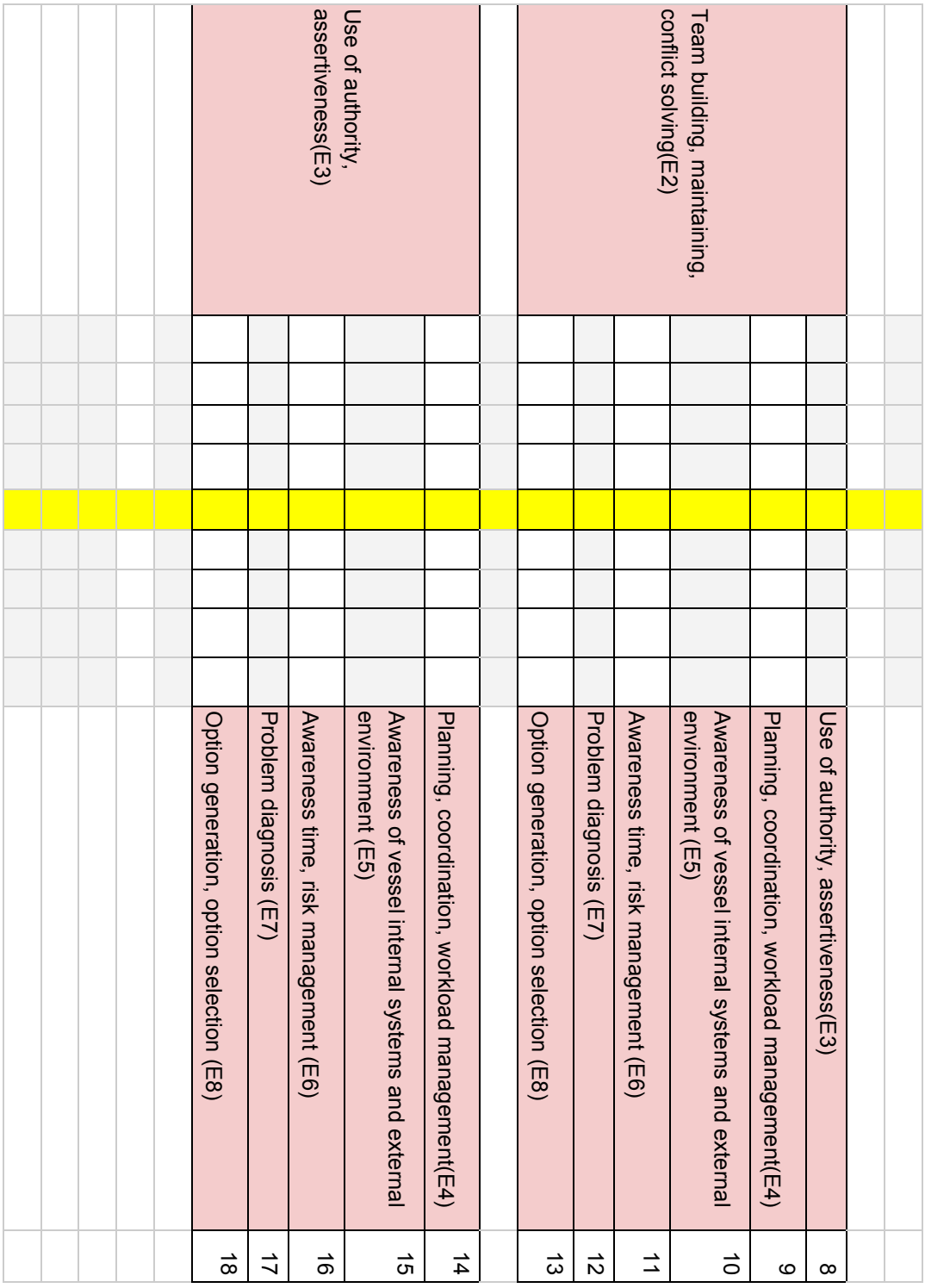
5 = Strongly important: Experience and judgment strongly favor Option X.

7 = Very strongly important: Option X is very strongly favored and it's dominance demonstrated in practice.

9 = Extremely important: The evidence favoring the highest possible order of affirmation for option X.

	Option 1	<-- 9	7	5	3	1	--> 3	5	7	9	Option 2	
											Team building, maintaining, conflict solving (E2)	1
											Use of authority, assertiveness (E3)	2
											Planning, coordination, workload management(E4)	3
	Consideration of others, Supporting Others: (E1)										Awareness of vessel internal systems and external environment (E5)	4
											Awareness time, risk management (E6)	5
											Problem diagnosis(E7)	6
											Option generation, option selection(E8)	7

Document 10: NTS Weights



Document 11: NTS Weights

Planning, coordination workload management (E4)										Awareness of vessel internal systems and external environment (E5)	19
										Awareness time, risk management (E6)	20
										Problem diagnosis (E7)	21
										Option generation, option selection (E8)	22
Awareness of vessel internal systems Awareness of external Environment (E5)										Awareness time, risk management (E6)	23
										Problem diagnosis (E7)	24
										Option generation, option selection (E8)	25
Awareness of time, Risk management (E6)										Problem definition and diagnosis (E7)	26
										Option generation, option selection (E8)	27

Document 12: NTS Weights

Problem definition and diagnosis (E7)										Option generation, option selection, Outcome review (E8)	28
Please answer the following questions.		years									
Total years of experience at sea.											
As a Captain											
As chief officer											

Document 13: NTS Weights

Behavioral markers and Exemplar Behaviours

Behavioral markers are defined as " Observable behaviors of teams or individuals, non-technical behaviors that contribute to superior or substandard performance within a work environment (for example, as contributing factors enhancing safety or in accidents and incidents in aviation) " (Klampfer, Flin, & Helmreich, 2001)

The following behavioral markers are derived from the aviation industry non-technical skills evaluation literature, for evaluation of Bridge Resource Management (BRM) course effectiveness.

Their applicability to the Norwegian maritime bridge officers was confirmed by a questioner submitted to a group of senior bridge officers (Captains and chief officers) with acceptable consensus level. However, the relative weight of the elements did not produce an acceptable level of consensus requiring a rerun of the survey. Below is given the skill elements and examplar behaviors for your reference.

Please refer to the behavioral elements and their examples below and make a pairwise comparison of each element.

Thank you for your participation.

Behavioral elements	Good practice examples:	Poor practice examples:
E1 consideration of others, Supporting Others:	Takes notice of suggestions; Takes conditionn of other crew members in to account, Gives personal feedback; Helps other crew members when in need; offers assistance.	Ignores suggestions; Does not take in to account condition of others; Hesitates to help in demanding situations.

Document 14: NTS Weights

E2 Team building, maintaining, conflict solving	Establishes atmosphere for open communication; Encourages inputs and feedback from others; Does not compete with Others Keeps calm in interpersonal conflicts; Suggests conflict solutions.	Blocks open communication; Keeps barriers between crewmembers; Competes with others; Hesitates to help in demanding situations
E3 Use of authority, assertiveness	Takes initiative to ensure crew involvement; Takes command if situation requires; advocates own position: Reflects on suggestions of others; Motivates crew by appreciation and coaches when necessary.	Hinders or withholds crew involvement; Passive, does not show initiative for decisions; own position not recognizable: Ignores suggestions of others; Does not show appreciation for the crew, coaches very little or too much.
E4 Planning, coordination workload management	Encourages crew participation in planning and task completion; Plan is clearly stated and confirmed; If change of plan is necessary crew is Consulted; Clearly stated goals boundaries for task completion.	Does not involve crew in planning; intentions neither stated nor confirmed; No consultation of crew when changing plan, or follows blindly; Unclear goals and boundaries.
	Appropriately distributes tasks, checks and corrects appropriately: Correct priority of tasks. Adequate time allotment to complete tasks; notifies signs of stress and fatigue.	Acts alone without other crew members involvement. Secondary tasks prioritized interfering with primary duties. Workload is increase due to inadequate planning; Ignores signs of stress and fatigue.

Document 15: NTS Weights

E5	Awareness of vessel internal systems Awareness of external Environment	Monitors and reports, changes in systems' states; Acknowledges entries and changes to systems: Collects information about environment (position, weather and traffic): Shares key information about environment with crew; Contacts outside resources when needed (to maintain situation awareness).	Does not ask for updates; Does not signal awareness of changing systems. Does not enquire about environmental changes; Does not comment on relevant environmental factors, or is surprised by them; Operates a 'closed shop':
E6	Awareness of time, Risk management	Discusses time constraints with crew; Discusses contingency strategies; Identifies possible future problems; Considers and shares estimated risk of alternative options; Talks about possible risks for action in terms of crew limits; Confirms and states selected option/ agreed action.	Does not set priorities regarding time limits; Does not discuss relationship between past events and present/future; Is surprised by outcomes of past events. Inadequate discussion of limiting factors with crew; No consideration of limiting factors: Does not inform crew of decision path being taken.
E7	Problem definition and diagnosis	Gathers information to identify problem; reviews causal factors with other crew Members	Nature of problem not stated or failure to diagnose; No discussion of probable causes

Document 16: NTS Weights

E8		
Option generation, option selection, Outcome review	States alternative options; asks crew members for options.	Does not search for information; Does not ask crew for alternatives.
	Confirms and states selected option/ agreed Action	Does not inform crew of decision path being taken.
	Checks outcome against plan	Fails to check selected outcome against goal

Document 17: Norsk senter for forskringsdata

NSD Personvern
02.10.2018 12:09

Det innsendte meldeskjemaet med referansekode 876672 er nå vurdert av NSD.

Følgende vurdering er gitt:

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD ENDRINGER
Dersom behandlingen av personopplysninger endrer seg, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. På våre nettsider informerer vi om hvilke endringer som må meldes. Vent på svar før endringer gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 30.11.2018.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD finner at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:
- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
De registrerte vil ha følgende rettigheter i prosjektet: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20). Rettighetene etter art. 15-20 gjelder så lenge den registrerte er mulig å identifisere i datamaterialet.
NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.
Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Zoho Survey er databehandler i prosjektet. NSD legger til grunn at behandlingen oppfyller kravene til bruk av databehandler, jf. art 28 og 29.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp behandlingen ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!
Kontaktperson hos NSD: Silje Fjelberg Opsvik
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Document 18: Consent

Are you interested in taking part in the research project

"Evaluation of Bridge Resource Management Training"?

This is an inquiry about participation in a research project where the main purpose is to devise a method to evaluate the effectiveness of Bridge resource management Training (BRM). In this letter, we will give you information about the purpose of the project and what your participation will involve.

Purpose of the project

Despite the fact that companies spend millions of kroner on Bridge Resource Management (BRM) training there is no easy way to evaluate its effectiveness. Without effective evaluation

1. *It is not possible to justify the time and resources invested in the training.*
2. *It is not possible to improve the training.*
   Our aim is to produce a method that would allow training institutions and companies in Norway to evaluate the training outcome by.
   a. *Identifying key behavioral markers of nontechnical-skills.*
   b. *Validate the skills applicability to Norwegian officers, by submitting a questionnaire survey to subject matter experts who are masters and chief officers.*

   This is a Master's Thesis project in Maritime Management.

Who is responsible for the research project?

*The University Of Southeast Norway (USN)*is the institution responsible for the project.

Why are you being asked to participate?

You are being asked to participate because you are considered a subject matter expert. It is believed that Masters and Chief officers would have the necessary experience and expertise to know which behaviors and non-technical skills are expected/necessary for a navigator to perform his/her duty to

1. *Avoid an emergency or crisis situation.*
2. *Trapp causes of an emergency or crisis situation.*
3. *Mitigate damage if an emergency or crisis situation occurs.*

The survey is sent to selected shipping companies to be forwarded to Masters and Chief officers serving onboard their ships.

What does participation involve for you?

- *The survey will be conducted through a website https://survey.zoho.com.*
- *Demographic information, age, sex and number of years of experience are considered important factors in forming an opinion. These opinion differences help to identify training needs and help to design better training.*
- *« If you chose to take part in the project, this will involve that you fill in an online survey. It will take approx. 45 minutes. The survey includes questions about your agreement or disagreement of non-technical skills requirement. Your answers will be recorded electronically»*

Participation is voluntary

Participation in the project is voluntary. If you chose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be made anonymous. There will be no negative consequences for you if you chose not to participate or later decide to withdraw.

Document 19: consent

Your personal privacy – how we will store and use your personal data.
Whenever one visits a website an ip-address of a visiting computer is visible to the server. In this survey the option to collect ip-address is disabled. However since ip-address is considered personal data, we are required to inform you that we will only use your personal data for the purpose(s) specified in this information letter. We will process your personal data confidentially and in accordance with data protection legislation (the General Data Protection Regulation and Personal Data Act).

- *The survey may be accessible to the supervisor and the faculty for evaluation purpose.*
- *The data is protected by password and two-factor authentication.*
- *No contact information will be collected or stored.*
- *No recognizable personal data will be published*

What will happen to your personal data at the end of the research project?
The project is scheduled to end 15th November 2018.

- *At the end of the survey, the account at the website will be closed and all data will be deleted as per the privacy policy of the website. https://www.zoho.com/privacy.html*
- *The downloaded results will be encrypted and stored with password access.*

Your rights
So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data

What gives us the right to process your personal data?
We will process your personal data based on your consent.

Based on an agreement with the University of Southeast Norway, NSD – The Norwegian Centre for Research Data AS has assessed that the processing of personal data in this project is in accordance with data protection legislation.

Where can I find out more?
If you have questions about the project or want to exercise your rights, contact:

- *University of Southeast Norway*
  - *Tegegne Tefera, e-post 012557@usn.no phone +47 95 76 71 54*
  - *Supervisor Førsteamanuensis Dr. Salman Nazir, e-post salman.nazir@usn.no.*
- NSD – The Norwegian Centre for Research Data AS, by email: (personverntjenester@nsd.no) or by telephone: +47 55 58 21 17.

Yours sincerely,

Student	Tegegne Tefera	Tegegne Tefera
Researcher/supervisor	Salman Nazir	salman nazir

Author: **Klaus D. Goepel**
http://bpmsg.com

Revisions

Date	Comment
3/5/2012	1st draft based on AHPcalc vers **27.8.11** (single input); allows input from max. 7 participants
11.12.2012 released	Algorithm of Power Method to solve Eigenvalue was modified (new sheet 8x8) **resulting in much higher accuracy**. By default 12 iterations
08.02.2013 released	final tests
2/19/2013	corrected formatting problem of weights in summary sheet
5/7/2013	Introduced weights for individual participants (weighted geometric mean) in sheet multInp Extend to **20 participants**
12/24/2013	For the 3 most inconsistent judgments **the ideal judgment** resulting in lowest inconsistency is displayed
5/9/2014	Change of check for convergence of power method
7/26/2014	corrected wrong ref in multInp (Matrix13)
4/9/2015	changed name of consol. Matrix from MatrixC to m_p0 CHOOSE in Summary sheet now p_sel+1 and m_p0 included (IF cond removed) changed the min of 3 to 2 criteria
6/7/2015	correction for 2 criteria: sheet 10x10 Cell M41 limit to 12 CGI in summary sheet to "n/a" for n=2, text (2 - 10)
5/4/2016	Corrected display of the selected scale in the summary sheet
5/4/2017	Corrected AHP consensus indicator (H gamma max)
8/22/2018	Balanced scale replaced by generalized balanced scale (Balanced-n), Added adaptive scale. Power method max no of iterations increased from 12 to 20.
9/15/2018	Inconsistency matrix pi/pj limited between 1/9 and 9 Error estimate for EV Error estimate for RGGM, Ordinal inconsistency Psi

Table 24
AHP tool credit

Document 20: BRM evaluation Norwegian maritime authority:(Sjøfartsdirektoratet, 2017)

Sjøfartsdirektoratet
Norwegian Maritime Authority

Bedømmelse av BRM-kompetanse

Etternavn, fornavn	Fødselsnummer (11 siffer)
Gateadresse	Postnummer
Poststed	Land
Stilling som:	Nasjonalitet

1. **Dokumentasjon**
   Assessor må ha assessorbevis utstedt fra Sjøfartsdirektoratet. Assessor må i tillegg ha dokumentasjon på tjeneste i stilling som offiser med minimum seks måneders fartstid etter 1. januar 2012 på skip med godkjent sikkerhetsstyringssystem i henhold til ISM-koden. Simulatorinstruktører og lærere som underviser i BRM-kompetanse i henhold til STCW-78-konvensjonen, med endringer, oppfyller kravet til godkjent assessor.
   Assessors bedømmelse skal baseres på krav i STCW-78-konvensjonen, med endringer, tabell A-II/1, kolonne 2 og 4. BRM-kompetanse må bedømmes gjennom observasjon, og må gjennomføres om bord eller ved bruk av brosimulator.
   Dette skjema for bedømmelse skal benyttes av assessor.

2. **Hvem som kan dokumentere BRM-kompetanse gjennom en bedømmelse av godkjent assessor:**
   Dekksoffiserer som har tjenestegjort i minimum seks måneders fartstid etter 1. januar 2012 på skip med godkjent sikkerhetsstyringssystem, oppfyller kravet til godkjent erfaring fra tjeneste. Disse kan da få sin kompetanse bedømt av en godkjent assessor.

3. **Bridge Resource Management kompetanse som skal vurderes**
   Det er kompetanse etter "*STCW Section A-II/1 function: " Navigering på det operative nivået, ledelse av ressurser på broen*" som skal bedømmes.

Document 21: BRM evaluation Norwegian maritime authority:(Sjøfartsdirektoratet, 2017)

Sjøfartsdirektoratet Bedømmelse av BRM-kompetanse

4. Assesors kriterier for evaluering av kompetanse
Vurdering av kompetanse skal skje under observasjon ombord, eller ved bruk av simulator. Under observasjon og vurdering skal assessor legge vekt på at:

- operasjonene er i samsvar med gjeldende regler og prosedyrer,
- at planleggingen ar hensyn til tilgjengelige ressurser og,
- at instrukser og kommunikasjonen er klar og entydig,
- at effektivt lederskap utvises
- alt involvert mannskap (team members) blir gjort forstått med skipets situasjon og operasjonell status,
- at beslutningene er effektive ut I fra gitte omstendigheter
- at beslutninger og operasjoner er bedømt som effektive og i samsvar med gjeldende regler.

Ledelse av ressurser på broen Kjennskap til prinsipper for ledelse av ressurser på broen herunder:	Kompetanse område; *STCW Section A-II/1 function: "Navigering på det operative nivået, ledelse av ressurser på broen"*	Assessment
Fordeling, tildeling og prioritering av ressurser	Ressursene er fordelt og tildelt etter behov i riktig prioritert rekkefølge for å utføre nødvendige oppgaver.	
Effektiv kommunikasjon	Kommunikasjon gis og mottas klart og utvetydig	
Bestemthet og lederskap	Effektiv lederskapsatferd identifiseres	
Innhenting og opprettholdelse av situasjonsforståelse	Tvilsomme beslutninger og/eller handlinger fører til egnet utfordring og reaksjon.	
Vurdering av gruppens erfaring	Medlemmer av gruppen deler presis forståelse av nåværende og forutsagt tilstand for fartøyet, navigeringsvei og ytre omstendigheter	

4. Attestering og godkjenning fra assessor
Assessor med assessorbevis utsted fra Sjøfartsdirektoratet skal verifisere at det ovenfor nevnte er observert og at angjeldende søker innehar nødvendig kompetanse.

Dato, sted	Assessor stempel
Signatur assessor	

Table 25
STCW Table A-V/2: (IMO, 1995b)

Table A-V/2
Specification of minimum standard of competence in crisis management and human behaviour

Column 1	Column 2	Column 3	Column 4
Competence	**Knowledge, understanding and proficiency**	**Methods for demonstrating competence**	**Criteria for evaluating competence**
Organize shipboard emergency procedures	Knowledge of: .1 the general design and layout of the ship .2 safety regulations .3 emergency plans and procedures The importance of the principles for the development of ship-specific emergency procedures, including: .1 the need for pre-planning and drills of shipboard emergency procedures .2 the need for all personnel to be aware of and adhere to pre-planned emergency procedures as carefully as possible in the event of an emergency situation	Assessment of evidence obtained from approved training, exercises with one or more prepared emergency plans and practical demonstration	The shipboard emergency procedures ensure a state of readiness to respond to emergency situations
Optimize the use of resources	Ability to optimize the use of resources, taking into account: .1 the possibility that resources available in an emergency may be limited .2 the need to make full use of personnel and equipment immediately available and, if necessary, to improvise Ability to organize realistic drills to maintain a state of readiness, taking into account lessons learnt from previous accidents involving passenger ships; debriefing after drills	Assessment of evidence obtained from approved training, practical demonstration and shipboard training and drills of emergency procedures	Contingency plans optimize the use of available resources Allocation of tasks and responsibilities reflects the known competence of individuals Roles and responsibilities of teams and individuals are clearly defined

Document 22: STCW: (IMO, 1995b)

Column 1	Column 2	Column 3	Column 4
Competence	**Knowledge, understanding and proficiency**	**Methods for demonstrating competence**	**Criteria for evaluating competence**
Control response to emergencies	Ability to make an initial assessment and provide an effective response to emergency situations in accordance with established emergency procedures *Leadership skills* Ability to lead and direct others in emergency situations, including the need: .1 to set an example during emergency situations .2 to focus decision making, given the need to act quickly in an emergency .3 to motivate, encourage and reassure passengers and other personnel *Stress handling* Ability to identify the development of symptoms of excessive personal stress and those of other members of the ship's emergency team Understanding that stress generated by emergency situations can affect the performance of individuals and their ability to act on instructions and follow procedures	Assessment of evidence obtained from approved training, practical demonstration and shipboard training and drills of emergency procedures	Procedures and actions are in accordance with established principles and plans for crisis management on board Objectives and strategy are appropriate to the nature of the emergency, take account of contingencies and make optimum use of available resources Actions of crew members contribute to maintaining order and control
Control passengers and other personnel during emergency situations	*Human behaviour and responses* Ability to control passengers and other personnel in emergency situations, including:	Assessment of evidence obtained from approved training, practical demonstration and shipboard training and drills of emergency procedures	Actions of crew members contribute to maintaining order and control

Document 23: STCW: (IMO, 1995b)

Column 1	Column 2	Column 3	Column 4
Competence	**Knowledge, understanding and proficiency**	**Methods for demonstrating competence**	**Criteria for evaluating competence**
Control passengers and other personnel during emergency situations (*continued*)	.1 awareness of the general reaction patterns of passengers and other personnel in emergency situations, including the possibility that: .1.1 generally it takes some time before people accept the fact that there is an emergency situation .1.2 some people may panic and not behave with a normal level of rationality, that their ability to comprehend may be impaired and they may not be as responsive to instructions as in non-emergency situations .2 awareness that passengers and other personnel may, *inter alia*: .2.1 start looking for relatives, friends and/or their belongings as a first reaction when something goes wrong .2.2 seek safety in their cabins or in other places on board where they think that they can escape danger .2.3 tend to move to the upper side when the ship is listing .3 appreciation of the possible problem of panic resulting from separating families		

Document 24: STCW: (IMO, 1995b)

Column 1	Column 2	Column 3	Column 4
Competence	**Knowledge, understanding and proficiency**	**Methods for demonstrating competence**	**Criteria for evaluating competence**
Establish and maintain effective communications	Ability to establish and maintain effective communications, including: .1 the importance of clear and concise instructions and reports .2 the need to encourage an exchange of information with, and feedback from, passengers and other personnel Ability to provide relevant information to passengers and other personnel during an emergency situation, to keep them apprised of the overall situation and to communicate any action required of them, taking into account: .1 the language or languages appropriate to the principal nationalities of passengers and other personnel carried on the particular route .2 the possible need to communicate during an emergency by some other means, such as by demonstration, or by hand signals or calling attention to the location of instructions, muster stations, life-saving devices or evacuation routes, when oral communication is impractical .3 the language in which emergency announcements may be broadcast during an emergency or drill to convey critical guidance to passengers and to facilitate crew members in assisting passengers	Assessment of evidence obtained from approved training, exercises and practical demonstration	Information from all available sources is obtained, evaluated and confirmed as quickly as possible and reviewed throughout the emergency Information given to individuals, emergency response teams and passengers is accurate, relevant and timely Information keeps passengers informed as to the nature of the emergency and the actions required of them

Glossary

ADDIE Analysis, Design, Develop, Implement, Evaluate. 15

AHP Analytical Hierarchy Process. 5, 27, 33, 35–37, 43

AIBN Accident Investigation Board Norway. 49

ANTS Anesthetists' Nontechnical Skills. 12, 28

AQP Advanced Qualification Program. 16

BMS Behavioural Marker Systems. 1, 2, 4–6, 10–13, 17, 21–23, 25–28, 31–33, 35–38, 41, 45–48, 50, 51

BRM Bridge Resource Management. 1, 2, 8–12, 28, 32, 37, 38, 46, 47, 49, 50

CRM Crew Resource Management. 2, 4, 8–10, 15–17, 20–23, 26–28, 38, 47, 48, 50

CVR Content Validity Ratio. 32, 42, 47

DAIBN Defense Accident Investigation Board Norway. 49

EMSA European Maritime Safety Agency. 8

ERM Engine Resource Management. 9, 49, 50

GIHRE Group Interaction in High Risk Environments. 4, 13, 25

IMO International Maritime Organization. 2, 9, 10, 50

JAR European Joint Aviation Requirements. 23

JARTEL Joint Aviation Requirements: Translation and Elaboration of Legislation. 24

LOFT Line Oriented Flight Training. 22

LOSA Line Operations Safety Audit. 6, 17, 23, 26, 27

MRM Maritime Resource Management. 2, 3, 12, 50

NANTS-no Nurse Anaesthetists' Non-Technical Skills-Norway. 12, 28

NASA/UT NASA and the University of Texas Human Factors Research Project. 13, 22

NOTECHS Non-Technical Skills assessment system. 4, 6, 13, 23–27, 38, 40

NTS Non-technical skills. 2, 3, 5, 10, 12, 13, 18, 21, 27, 35, 36, 46, 47

RSSB Rail Safety and Standards Board. 27

SA Situational Awareness. 47

STCW Standards of Training, Certification and Watch keeping for Seafarers. 9, 10

TEI The Training Evaluation Inventory. 19

TEM Thereat and Error Management. 17, 23

USN University of South-East Norway. 12, 47, 50