  DOI: 10.21339/2464-353x.3.2.164

Can you teach an old seadog new tricks? Experimental evaluation of BRM training in the commercial fleet

Sturle Danielsen Tvedt, Western Norway University of Applied Sciences, Simsea Real Operations | sturle.tvedt@gmail.com
Roar Espevik, Royal Norwegian Naval Academy
Helle Asgjerd Oltedal, Western Norway University of Applied Sciences
Guro Persdotter Fjeld, Western Norway University of Applied Sciences
Frode Voll Mjelde, Royal Norwegian Naval Academy

Abstract

Objective: The objective of the present study was to evaluate the effectiveness of Crew Resource Management (CRM) training in the commercial shipping fleet – termed Bridge Resource Management (BRM) training.

Background: CRM training has been widely employed and researched in several high reliability settings. However, there is a lack of experimental studies assessing CRM training in commercial shipping.

Method: An experimental pretest – posttest study measuring satisfaction with training, knowledge, attitudes, and team behavior in bridge simulators. Five hypotheses were made; H1) The BRM training will receive positive evaluation, H2) BRM training will improve knowledge, H3) BRM training will improve attitudes, H4) BRM training will improve behavior, H5) The relationship between Teamwork and Mission success is positively mediated by Situation awareness.

Results: H1 was fully supported. H2 was fully supported. H3 was partly supported. H4 was not supported. H5 was fully supported.

Conclusion: The training was positively evaluated and improved knowledge and some of the targeted attitudes. Behavior could not be shown to improve with statistical significance, but it cannot be ruled out that a stronger experimental design and increased sample size would yield significant results. Relations among behavior measures confirms established CRM theory.

Application: The present study provides supporting evidence that BRM training can indeed improve safety-relevant knowledge and attitudes. However, to improve behavior on the bridge, training should be adapted to specific work procedures.

Introduction

The International maritime organization (IMO) has declared shipping as one of the most dangerous industries in the world (2011). Personnel on board face complex and dangerous machinery, often made worse by heavy sea and challenging navigation. Human errors in this environment can have serious consequences, leading to collisions and explosions with a potential to be both enormous and tragic. The risk becomes even greater when operating far away from first-responder assistance (Hetherington, Flin, Mearns, 2006). To prepare for such a high-risk environment and to avoid accidents, proper training becomes a major concern for maritime organizations (International Maritime Organization, 2002).

Crew Resource Management (CRM) training emerged in the airline industry after several major accidents during the 1970s, when the authorities acknowledged that technical competence alone was not sufficient to guarantee safe performance (Rutherford, Flin & Mitchell, 2012). 30 years later the CRM training has transferred to other high-risk organizations, such as healthcare, military, offshore industry and nuclear plants. In 2010 CRM training became mandatory for the maritime industry in the Manila amendments to the Standards of Training, Certification and Watchkeeping for Seafarers (STCW) regulations (International Maritime Organization, 2011).

However, CRM training is not unproblematic. Despite decades of CRM training of teams in high-risk organizations such as shipping, health care and military there is still considerable uncertainty whether this type of training actually increases safety (Salas et. al., 2006). One reason could be that there seems to be no established agreement as to what CRM training should entail (O'Connor, 2008), making it problematic to transfer training strategies between domains. Musson and Helmreich (2004) underline that CRM training tends to be domain specific, both organizationally and culturally. Specifically, the automatic transplantation of CRM training in aviation to maritime BRM training has been suggested as an explanation for the lack of results within bridge crews (O'Connor, 2011). In the maritime domain, the regulations put forward in 2010 (International Maritime Organization, 2011) makes BRM and ERM training mandatory to achieve maritime certificates. IMO does not however, present or list specific contents for BRM courses in a separate paragraph in the STCW manual. Typical CRM topics are spread out and inserted into various chapters and paragraphs that existed before the 2010 amendments, which make it difficult for instructors and training establishments to construct standardized BRM curriculums.

Despite the domain specific differences, the objective of resource management courses is established through a common conceptual perspective. Salas et. al. (2006) stated in a review that CRM courses (i.e. BRM or ERM in the maritime domain) are intended to increase knowledge, awareness and skills around the importance of clarity of roles, clear communication, and situation awareness. Within this perspective, highly reliable organizations will work towards avoiding accidents by collectively identifying and managing evolving threats. Hence, team members are encouraged to continuously scan for threats and to speak up when they identify potential threats, regardless of their status in the hierarchy or their defined role (Weick, 2002). A CRM course is intended to train and build awareness around teamwork behavior that enhance a common understanding of the situation at hand, eventually resulting in higher performance or mission success. Theoretical perspectives involve individual learning (mental models), teamwork behaviors that enhance and maintain shared mental models and situation awareness (Endsley, 2015, Espevik & Olsen, 2013).

Previous studies have predominantly used undergraduate students (e.g. Stout 1999) or aviator trainees (Salas, E., Cannon-Bowers, Rhodenizer, & Bowers, 1999). Salas, DiazGranados, Weaver, & King (2008) underlined the importance of studying teams in the wild (i.e. when operating their normal job at sea). In the maritime domain, there is a lack of true experimental studies evaluating BRM courses, especially for the commercial fleet. The few existing studies of BRM and related training is either limited to participant satisfaction (Håvold et al., 2015), or to navy samples (O'Connor, 2011; Röttger, 2016).

The present study will examine the effectiveness of BRM training on knowledge, attitudes and behavior in experienced bridge officers through an experimental evaluation of a commercial BRM training program performed for a Norwegian ship owner. The core of this BRM training program evolves around teamwork behavior that builds situation awareness and shared mental models.

Theory

The present theoretical perspective is outlined as an Input-Process-Outcome model (IPO; see Figure 1). The BRM course was constructed and performed to give knowledge about, attitudes towards, and training of, teamwork behavior. This teamwork behavior is understood as Input, which will result in a Process, i.e. better mental models, which enhance Situation awareness (understanding - "what is happening") and Shared mental models (coordination behavior - "what to do"). The outcome of all this is better performance (e.g., mission success).

The input in the BRM course is theoretical and practical training on teamwork behavior. Salas, Sims and Burke (2005) presented a model of teamwork integrating

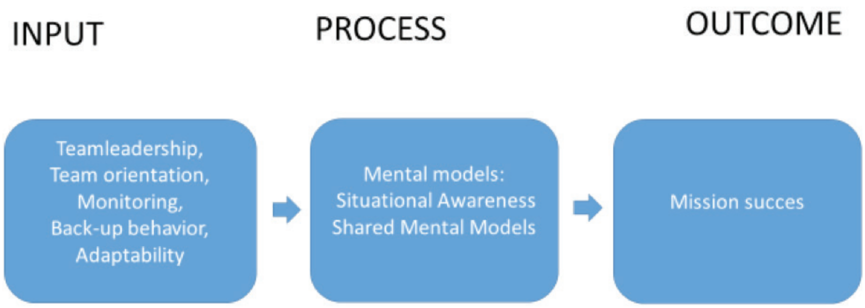

Figure 1. An Input-Process-Outcome model of BRM training.

the most commonly discussed variables that had the greatest effect on team performance. They suggested 5 core components, all of them vital to develop and flourish situational awareness and shared mental models. The first core component is *team leadership*, which entails the ability to direct and coordinate the activities of other team members. Second, *mutual performance monitoring*, which is the ability to apply appropriate task strategies to develop common understanding of the team environment. Third, *backup behavior,* which entails team members' ability to anticipate each other's needs through knowledge about their responsibilities. Fourth, *adaptability,* which concerns the team's ability to adjust team strategies and alter course of action based on information gathered from the environment. The last one, *team orientation* is an attitude characterized by a tendency to take team members behavior and input into account during group interaction, and that team goals are placed above individual goals.

Knowledge about, attitudes towards and training of these teamwork behaviors intends to enhance a team member's mental models of the environment he or she is supposed to operate in. Mental models are the mechanisms whereby humans are able to generate descriptions of a system's purpose and form, explanations of its functioning and observed states, and predictions of future system states (Rouse & Morris, 1986). This is in line with Endsley (2015), as the three stages of Situation awareness (SA) correspond to the three purposes of mental models, namely detecting (elements in the environment), explaining (their meaning) and predicting (their future status).

Endsley (1995) defines situation awareness as 'The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future'. This indicates that SA is as dynamic as the situation. When new information emerges, SA needs to be updated and changed, thus made possible by the five core teamwork behavior (Salas et.al., 2005). Although Endsley labels SA as a state, it can be argued that she describes three cognitive processes or functions, namely perception, comprehension and projection (Salmon et al., 2006). Perception involves detection of critical signals that are clearly observable and meaningful pieces of information (Level 1). Comprehension (Level 2) involves interpreting and combining relevant perceived information in order to grasp a correct understanding related to a goal. Projections (Level 3) represent forethought where the decision makers predict the status in near future. Thus, SA could be considered as the decision maker's internal model and forms the basis for decisions made. Lack of SA has been stated as the cause of human errors in critical situations. Sneddon, Mearns and Flin (2006) reported that 67% of errors made in their material were caused by lack of Level 1 SA; a failure to detect critical signals. 20 % of the accidents were caused by a lack of correct understanding of the situation (Level 2) and 13% was caused by a failure to correctly predict possible future states of the situation (Level 3). When further investigating the causes of poor SA, the main reason for a failure in Level 1 was found as failure to monitor or distraction causing reduced attention to the task at hand. In addition, a low ability to prioritize the available information resulted in information overload, and critical signals were not detected.

Operational decision making is most often done in a team setting. A related concept to SA is "shared mental models" (SMM). SMM is defined as a shared organized understanding and mental representation of key elements of the team's relevant environment. These shared mental models enable team members to form accurate explanations and expectations of the task. This will in turn enable team members to coordinate their actions and adapt their behavior to the demands of the task and to fellow team members (Converse, Cannon-

Bowers Salas, 1993). SMMs are assumed to enable team members to predict task needs and the actions of other team members, and thus enable them to adapt their own behavior accordingly without communicating explicitly. In a meta-analysis, DeChurch & Mesmer-Magnus (2010) showed a number of studies indicating that SMMs contribute to increased team effectiveness (e.g. Stout, Cannon-Bowers, Salas, & Milanovich, 1999). In a series of simulator studies Espevik, Johnsen, Eid & Thayer (2006) found that operational submarine attack teams with a high degree of SMMs had better performance than other teams. Furthermore, Naval teams with high degree of SMMs showed better coordination and performance when two teams had to coordinate their effort towards a common goal, compared to low SMM-teams (Espevik, Johnsen & Eid, 2011 a). Finally, high SMM teams showed improved learning using cross training in high intensity simulation compared to low SMM-teams (Espevik, Johnsen & Eid, 2011 b). Thus, training focusing on the importance of developing SMMs is important within the maritime domain.

As previously described, CRM training started in the aviation community and quickly spread to other high reliability settings such as armed forces, nuclear energy and healthcare. Since the 1990s healthcare has been the leading domain for research, providing ample evidence for the effectiveness of CRM training (Flin, O'Connor, & Crichton, 2008; Hughes et al., 2016; Weaver et al., 2010). However, when it comes to the maritime domain, studies are scarce. Most of them are set within the surface warfare community (O'Connor, 2011; Röttger et al., 2013; 2016) and the one set in the commercial fleet only deals with the self-reported post training satisfaction (Håvold et al., 2015). Hence, the present study sets out to contribute to the knowledgebase regarding BRM courses for the commercial fleet by evaluating the effectiveness of such a course arranged for a Norwegian ship owner.

Evaluating BRM effectiveness

Kirkpatrick's (1976; 2009) hierarchy is an often-used and valued framework for guiding training evaluation, and consists of four different levels: reactions, learning, behavior and organizational impact (see O'Connor, Campbell, Newon, Melton, Salas and Wilson, 2008; Salas et.al., 2006). This framework entails firstly, *reaction* to cover the degree to which the participants find the time spent worthwhile, or put simply, if they like the training. Secondly there is the level of *learning* which means that the training was understood and absorbed. Learning consists of acquiring knowledge, and personal knowledge is defined as 'the cognitive resources which a person brings to a situation that enable him or her to think and perform' (Eraut, 2000). In assessment terms, learning corresponds to written tests assessing theoretical knowledge. Thirdly, *behavior* is the assessment of whether knowledge learned in training actually transfers to behaviors in a work setting (in our case a work setting in a simulated environment). Fourthly the highest level, *Organizational impact*, to provide evidence for improved safety and effectives in the daily operation of the organization. The findings on the three last levels, learning, behavior and organizational impact are scarce (Salas et.al., 2006), and too our knowledge almost non-existent within the commercial maritime domain.

Hypotheses

There is considerable evidence for self-reports that participants value CRM training (DeChurch, et.al. 2010; Espevik, Saus, & Olsen, 2017) or level one, reaction in the hierarchy proposed by Kirkpatrick (2009). Although it is a very basic indicator of training quality, it is still a necessary part of training, and with the ample evidence in the literature a mean evaluation score higher than "uncertain" is expected.

H1) The BRM training should receive positive evaluation.

Although the few studies most relevant produces mixed evidence for increasing knowledge (O'Connor, 2011; Röttger et al., 2016), the training program evaluated in the present study was informed by these findings. Specifically, O'Connor (2011) presented failing to adapt training to the maritime domain as the cause of his null-finding regarding knowledge. This was a key instructive finding informing the establishment of the present training program, and great care was taken to adapt the training to the maritime domain. In addition, improving knowledge is a key outcome for the present training. Hence, it is reasonable to assume that the BRM training should in fact improve knowledge.

H2) BRM training will improve knowledge.

Both of the most relevant studies showed no evidence for improved attitudes (O'Connor, 2011; Röttger et al., 2016). However, again, O'Connor (2011) failed to adapt training to the maritime domain, and Röttger et al. (2016) only used classroom training. In addition, improving attitudes is a key learning objective for the present training. Hence, we expect the training to improve attitudes.

H3) BRM training will improve attitudes.

No study has yet produced evidence for behavior change by BRM training in the maritime domain. Furthermore, fundamental tenets of the cognitive psychological paradigm would suggest that knowledge and attitudes are indirect determiners of behavior (Eagly & Chaiken, 1993). However, it has long been an established fact that change in behavior is more difficult than change in knowledge (Hollenbeck, van Knippenberg, & Ilgen, 2017), and this challenge is confounded when applied to a group setting such as in the present study (Lewin, 1943; 1947). Nevertheless, the Kirkpatrick paradigm suggests behavior change as a separate level of evaluation. Furthermore, in the present study, it is of course a main motivation for BRM courses that training teamwork behavior should lead to behavioral change.

H4) BRM-training will improve behavior

According to established theory on the role of Teamwork and situation awareness for Mission success, we expect these behavioral measures to be related at Time 1 (prior to training) and at Time 2 (subsequent to training). In other words: the different behavioral measures are related for each team in each exercise, such that a there is a relationship between a team's Teamwork to their Mission success positively mediated by Situation awareness (see e.g. Espevik et al., 2006; 2011a; 2011b; Stout et al., 1999). See Figure 2. for a conceptual model for H5.

H5) The relationship between Teamwork and Mission success is positively mediated by Situation awareness at each point in time.

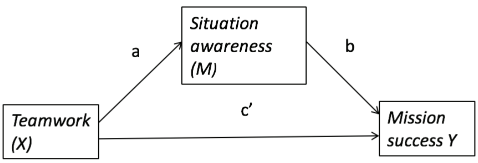

Figure 2. Conceptual model of the relationship of Teamwork mediated through Situation awareness on Mission Success. Conceptual model for the mediation models is showed in the results, tables 5 and 6. Constant coefficients, denoted i_M and i_Y in the table are not represented in the figure as they have only technical statistical interest.

METHODS

Design

The study was set up as a pretest - posttest experimental design. All participants were subject to surveys and simulator tests prior to, and subsequent to, BRM training (see Table 1). All participants received the same theoretical training and practical simulator exercises. Control was achieved with random assignment of the two scenarios (A and B) used for simulation exercises so that one half of participants did A-B and the other half did B-A. Hence, all participants received treatment between pre- and posttest. In addition, random assignment was employed of teams to bridge simulator and assessor for each course week comprising four teams per week. Thus, any variation in execution between course weeks affected each experimental condition equally. Assignment of participants followed a random stratified approach, stratifying according to rank. The result was that most teams had one team member with superior rank to the rest, usually a captain. No teams had junior officers only. Each team consisted of 3 members, except for a minor number of teams with only two participants due to absentees.

Training and Simulator Scenarios

The BRM course consisted of four simulator scenarios in total. The first is a simple transit scenario without critical incidents to familiarize the participants with the simulator. The third scenario was a Search and Rescue (SAR) scenario given the same to all groups as part of the training. The second (scenario A) and fourth (scenario B) scenarios presented were the two scenarios developed specially for the present study, although they could have been used in any regular BRM course: Scenario 2 (A) involved personnel injury on deck in an offshore setting, meaning that the bridge team must make the transition from a standard cargo delivery operation to a critical evacuation of injured personnel, a taxing team effort inducing stress. Scenario 4 (B) involved GPS deviance during a navigational exercise in coastal waters, testing the team's ability to maintain Situation awareness; discovering the GPS deviance through active information gathering and understanding that the GPS derived position is incorrect in the electronic charts (Electronic Chart and Display and Information System – ECDIS).

The total length of the BRM training was five days. According to the Norwegian Maritime Authority's (NMA) interpretation at the time of the international regulations concerning BRM training – Standards of Training, Certification and Watchkeeping for Seafarers (STCW) (International Maritime Organization, 2011), the recommended length of a BRM course was four days. The study included the recommended four days for training, consisting of two days classroom lectures and two days simulation training and observation including feedback sessions – using scenario 3. The pre- and posttests constituted one extra day in duration, bringing this specific course up to five days.

Table 1. Schematic illustration of experimental design.

Randomly assigned Group	Pretest			Training		Posttest	
	Survey 1	Simulator exercise 1[a)]	Simulator exercise 2	Theoretical training	Simulator exercise 3	Simulator exercise 4	Survey 2
1	Knowledge Attitudes	Familari-zation scenario	Scenario A «Stress»	Classroom lecture and discussion	«Search & rescue»	Scenario B «Situation awareness»	Satisfaction Knowledge Attitudes
2	Knowledge Attitudes	Familari-zation scenario	Scenario B «Situation awareness»	Classroom lecture and discussion	«Search & rescue»	Scenario A «Stress»	Satisfaction Knowledge Attitudes

[a)] The familiarization is not a measurement and not the intended experimental manipulation but a necessary preparation of the participants to the equipment on the simulator bridge.

Participants

A total of 94 experienced bridge officers participated in the study, 95.7% men. The age span was 23 to 62, (M=39.9, SD=10.1) and the seniority span was .5 to 34 years (M=6.2, SD=7.2). 35.5% were ranked as captains, 23.7% Chief mates, and 40,9% were first mates. Due to some missing values in either pre- or post-forms, the questionnaire analyses were performed with 79 to 80 participants.

Instruments

Satisfaction: An 8-item rating scale was used as part of the training centre's standard evaluation form. The scale is developed to rate satisfaction with simulation training and contain the following statements: 1) The course was adequate to my previous knowledge. 2) The content held a high professional level. 3) The content was relevant for my work. 4) The simulators are realistic. 5) The simulator exercises are realistic. 6) Our teaching methods are good. 7) Course facilities are satisfactory. 8) All together a good course. A sum score was then computed based on the responses, and the scale had a satisfactory consistency (cronbach's alpha = .78).

Knowledge: A 10-item multiple-choice test was taken with permission from O'Connor (2011) and is described there. A sum score was then computed based on the responses.

Attitudes: A 26 item Ship Management Attiudes Questionnaire (SMAQ) was taken with permission from Andersen, Garay and Itoh (1997). Each statement was followed by a 5-point Likert scale (1 = Fully disagree; 5 = Fully agree). An explorative factor analysis (see Table 2) produced five attitudinal categories: 1) My stress, 2) Stress of others, 3) Communication and coordination, 4) Command and responsibility, and 5) Revealing short comings.

Psychometric analysis

Several variations of SMAQ questionnaires have followed from the original classification of Helmreich and Merrit (1998). However, they do not agree on a common factor structure, which is to be expected from the little extent of actual overlap in items. Hence, the present study needed to establish its own factor structure (see, Andersen et al., 1999; O'Connor, 2011; Röttger et al, 2012). A principal Axis Factor (PAF) with Oblimin (oblique) rotation was run on the SMAQ items at Time 1 and yielded an adequate factor solution explaining 63.6% of the variance. An examination of the s Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was favorable (KMO=.675), and Bartlett's test of Sphericity was significant. Eigenvalue's and the scree plot both suggested five factors (see Table 2.)

The are some correlations among the factors both at Time 1 and Time 2, indicating support for an oblique rotation (see Table 3.). In terms of test-retest reliability, the factors correlate around $r=.60^{***}$ with the exception of factor 3 (Stress of Others), which has a r of $=.27^{*}$.

In terms of internal consistency, the factors show acceptable values of Cronbach alpha for the first two factors, and increasingly lower factors for the other three, but here the alphas should be considered in light of the low number of items (see Table 3). There is also a decreasing trend from Time 1 to Time 2 in internal consistency, especially with factor 3. This is in some degree expected as the study actively aims to change the attitudes.

Behavior: The behavior was rated by three subject matter experts based on video-recordings of Scenarios A and B. These observers were blind to whether the exercises were performed pre or post training. Three measures of behavior were rated: Teamwork, Situation awareness and Mission success.

Table 2. Factor analysis of SMAQ with factor loadings Time 1.

Factors and items	Factor loadings F1	F2	F3	F4	F5
F1 Communication / coordination					
10. A debriefing and critique of procedures and decisions after critical situations is an important part of safety.	.735				
26. Seniors should delegate responsibilities to junior crews as parts of their training.	.698				
20. Crewmembers should monitor each other for signs of stress or fatigue.	.671				
16. When I detect an error I speak up.	.611				
F2 My stress					
14. I am more likely to make errors in an emergency.		.811			
2. Even when fatigued, I perform effectively during critical times of operation.		-.660			
18. I am less effective when stressed or fatigued		.533			
9. My decision-making ability is as good in emergencies as in routine conditions.		-.449			
F3 Stress of others					
22. Effective team co-ordination requires team members to take into account the personalities of the others participants.			.687		
19. My performance is not adversely affected by working with inexperienced crewmembers			-.539		
4. People should be aware of and sensitive to the personal problems of other crewmembers.			.438		
F4 Command and response					
12. Junior crewmembers should not question their senior officer's decisions.				.558	
1. Senior officers should encourage crewmember questions during normal operations and in emergencies.				-.558	
F5Revaling short comings					
17. I am ashamed when I make a mistake in front of other crewmembers.					.608
3. Asking for assistance makes one appear incompetent.					.511

Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. Factors: 1 = Communication /coordination; 2 = My stress; 3 = Stress of others; 4 = Command and responsibility; 5 = Revealing short comings. (Factor loadings =>.40).

Table 3. Correlation matrix of SMAQ factors at Time 1 and Time 2 with internal consistency [(a)].

		Factors Time 1					Factors Time 2.				
		1	2	3	4	5	1	2	3	4	5
Factors Time 1	1	(.71)									
	2	-.081	(.70)								
	3	-.185	-.104	(.50)							
	4	.111	.055	-.079	(.30[b])						
	5	-.254*	.215*	.093	.024	(.29[b])					
Factors Time 2	1	.681***	.119	-.306**	.352**	-.067	(.62)				
	2	.038	.677***	-.157	.124	.240*	.258*	(.55)			
	3	-.331**	-.109	.271*	-.022	-.034	-.406***	-.174	(.15)		
	4	.281*	.249*	-.341**	.591***	-.122	.507***	.322**	-.302**	(.32[b])	
	5	-.154	-.021	.094	-.187	.592***	-.164	.151	.046	-.115	(.23[b])

[(a)] Cronbach alphas on the diagonal. [(b)] Corrected Item-total correlation. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$ *(one-tailed)*

Teamwork: to extract teamwork, 8 teamwork constructs of the Royal Norwegian Naval Acadamy (RNoNA) assessment tool was used. The RNoNA assessment tool is designed to assess the performance of military teams participating in complex military training exercises (Mjelde et.al, 2016). The purpose is to evaluate the teams' ability to communicate critical information to maximize collective performance, based on Salas et.al (2006) five teamwork processes. For example, for Backup behavior, Subject matter experts rated the backup behavior they observed by rating the following claim on a 7-point Likert scale from strongly disagree (1) to strongly agree (7): The team showed a high degree of backup behavior, i.e. team members helped/assisted without being asked, or pushed information.

Situation awareness: Two clearly discernable incidents were identified within each scenario. For example, in the "Stress scenario" the first incident was the initial awareness regarding an injury on deck, and the second incident was the subsequent awareness when the situation had evolved to the point when it became clear that helicopter evacuation was 45 minutes away. These incidents were rated by subject matter experts in accordance with Endsley's (1995; 2015) description of levels of situation awareness. For example, perception (quickly perceived the injury), understanding (correctly understood the situation) and prediction (correctly predicted potential outcomes). A mean Situation awareness score was computed from the three observed levels of Situation awareness rated for each incident (six observations in total for each scenario).

Mission success: The same two incidents were also rated as to what degree the teams performed a satisfactory action to control the incident from 1 (strongly disagree) to 7 (strongly agree), in accordance with recommendation of Salas, Tannenbaum, Kraiger, and Smith-Jentsch (2012). A mean score for Mission success was computed from the two incident scores.

Procedure

A questionnaire containing instruments measuring knowledge and attitudes were filled out by the participants prior to- and following the training. The behavior measures were taken during two different scenarios pre- and post-treatment. After being introduced to the course and familiarized with the simulators, the pre-scenario was performed. Following classroom training and a simulator exercise with plenary feedback, the post-scenario was performed.

Participation was based on informed consent and the gathering of data was approved by the Norwegian Data Authority for Social Sciences (NSD).

Table 4. Satisfaction, attitudes, knowledge, and behavior in simulator pre- and post- training with t-tests for change and Cohen's d with 95% Confidence Intervals.

Variable	Time	M	N	SD	t	df	p (one-tailed)	d	95% CI of d
Satisfaction	Post	4.75	80	.548	12.312	79	.000	1.38	[1.08, 1.67]
Knowledge	Pre	5.5250	80	1.79292	-4.368	79	.000	-0.49	[-0.72, -0.26]
	Post	6.5125	80	1.92251					
Communication and coordination	Pre	4.5156	80	.49418	-1.558	79	.062	-0.17	[-0.39, 0.05]
	Post	4.5813	80	.44255					
My stress	Pre	3.3133	79	.78596	-2.388	78	.019	-0.27	[-0.49, -0.04]
	Post	3.4736	79	.67970					
Stress in others	Pre	3.6646	79	.71735	-1.878	78	.032	-0.21	[-0.43, 0.01]
	Post	3.7932	79	.61395					
Command and response	Pre	1.7278	79	.64944	.067	78	.473	0.01	[-0.21, 0.23]
	Post	1.7215	79	.72831					
Revealing short comings	Pre	2.0250	80	.86383	1.027	79	.155	0.11	[-0.11, 0.33]
	Post	1.9375	80	.82052					
Teamwork[a)]	Pre	5.4341	26	1.13751	.045	25	.483	0.09	[-0.30, 0.47]
	Post	5.4220	26	1.24298					
Situation awareness [a)]	Pre	4.5313	24	1.59309	-1.234	23	.118	-0.25	[-0.66, 0.16]
	Post	5.0451	24	1.67714					
Mission success [a)]	Pre	4.6944	24	1.71940	-.288	23	.388	-0.06	[-0.46, 0.34]
	Post	4.8333	24	1.96942					

[a)] *Measures for Teamwork, Situation awareness and Mission Success are given as team values only.*

Hypothesis testing Analyses

The basic analyses, including descriptive, correlations, t-tests, and factor analysis were performed using IBM SPSS version 24.0® for Windows 10®. Effect sizes and 95% confidence intervals for these were computed using the MBESS package, version 3.5.1, with MBESS package version 4.4.3, (Kelley, 2018). H5 was tested with the process macros developed by Hayes (2013) through IBM SPSS 24.bbb0. The macros are based on standard ordinary least squares (OLS) regression (see Figure 2 for a conceptual model). As demonstrated by Preacher and Hayes (2004), this macro produces a test that is more rigorous than that of Baron and Kenny (1986) and at the same time avoids the bias of the Sobel (1982) approach. Preacher and Hayes (2004) achieved this by employing a bootstrapping procedure. Bootstrapping works by basing inferential procedures on concrete sampling distribution from the sample at hand, rather than traditional sampling distribution created by a hypothetical infinite number of samples from the population of interest (Efron, 1982). The concrete sampling distribution thus reflects the distribution of the sample, rendering the assumption of normality superfluous, and allows the testing of mediators on small samples (Preacher & Hayes, 2008). A bootstrap sample of 10,000 was drawn (with replacement) and used for analysis of the mediation model.

Results

Descriptive results are given in Table 4 presenting measures pre- and post-training. The results of the hypothesis testing can be found in tables 4, 5 and 6 and is commented according to the order of the hypotheses given in the introduction.

Table 5. Regression results for the Teamwork mediation model at Time 1 (prior to manipulation) with results for alternative models. Unstandardized OLS Regression Coefficients with Confidence Intervals (Standard Errors in parentheses) Estimating Situation awareness and Mission Success.

		Situation awareness (*M*)			Mission Success (*Y*)	
		Coeff.	95% CI		Coeff.	95% CI
Teamwork (*X*)	$a_1 \rightarrow$	0.736** (0.241)	0.237, 1.235	$c' \rightarrow$	0.342 (0.200)	-0.072, 0.757
Situation awareness (*M*)				$b \rightarrow$	0.812*** (0.146)	0.510, 1.115
Constant	$i_M \rightarrow$	0.925 (1.352)	-1.872, 3.721	$i_Y \rightarrow$	-1.052 (0.954)	-3.032, 0.927
Model Summary	R^2 = .288, F (1, 23) = 9,312**			R^2 = .739, F (2, 22) = 31.082***		

Bootstrap result for indirect effects				
Indirect effect	*M*	*SE*	LL 95%	UL 95%
Hypothesized Model	0.5981**	0.2262	0.3083	1.1489
Alternative Model 1[d)]	0.1339[ns]	0.0933	-0.0232	0.5304
Alternative Model 2[e)]	0.0493[ns]	0.1700	-0.5668	0.3846
Alternative Model 3[f)]	0.3251[ns]	0.1966	-0.0774	1.0064

**p < .05 **p < .01 ***p < .001. a = the direct effect of X on M. i_M = the direct effect of the constant on M. c' = the direct effect of X on Y. b = the direct effect of M on Y. i_Y = the direct effect of the constant on Y. [d)] Situation awareness on Mission Success mediated by Teamwork. [e)] Mission success on Teamwork mediated by Situation awareness. [f)] Mission Success on Situation awareness mediated by Teamwork.*

H1 posited that the BRM course would be evaluated higher than "uncertain-positive". As illustrated in Table 4, the difference between the mean score and "uncertain-positive" is statistically significant and a large effect.

H2 posited that the BRM course would increase the knowledge. As illustrated in Table 4, the increase in Knowledge is statistically significant and a medium effect.

H3 posited that the BRM course would improve attitudes. As illustrated in Table 4, there is a statistically significant improvement for Stress awareness, Team consideration, and very close to an improvement for Communication and coordination. These are all small effects, and the confidence interval for Communication and coordination includes 0, indicating a likelihood of no effect. The attitudes for Authoritarianism and Weakness toleration, however, showed no statistically significant change. In terms of effect sizes, they are also small to negligible for both of these, the confidence intervals include 0.

H4 posited that the BRM course would improve Behavior. However, there are no statistically significant changes in the behavioral measures. In terms of effect sizes, they are also small to negligible and for all of these, the confidence intervals include 0.

H5 posited that the relationship between Teamwork and Mission success is positively mediated by Situation awareness at each point in time. According to Tables 5 and 6, the hypothesized mediation model was significant at both points in time at alpha level .01 or higher (pre- and post-training). In addition, none of the other possible mediation models were statistically significant.

Table 6. Regression results for the Teamwork mediation model at Time 2 (subsequent to manipulation) with results for alternative models. Unstandardized OLS Regression Coefficients with Confidence Intervals (Standard Errors in parentheses) Estimating Situation awareness and Mission Success.

		Situation awareness (*M*)			Mission Success (*Y*)	
		Coeff.	95% CI		Coeff.	95% CI
Teamwork (*X*)	a →	0.909*** (0.221)	0.454, 1.365	c′→	0.113 (0.181)	-0.262, 0.488
Situation awareness (*M*)				b →	0.886*** (0.128)	0.620, 1.151
Constant	i_M →	-0.364 (1.241)	-2.926, 2.198	i_Y →	0.016 (0.782)	-1.601, 1.633
Model Summary	R2 = .414, F (1, 24) = 16.971***			R2 = .799, F (2, 23) = 45.722***		

Bootstrap result for indirect effects				
	M	*SE*	LL 95%	UL 95%
Indirect effect	0.8050***	0.2052	0.4949	1.3783
Alternative Model 1[d)]	0.0514[ns]	0.0859	-0.0852	0.3446
Alternative Model 2[e)]	0.2700[ns]	0.2134	-0.1534	0.5344
Alternative model 3[f)]	0.1377[ns]	0.2227	-0.2136	0.5877

**p < .05 **p < .01 ***p < .001. a = the direct effect of X on M. i_M = the direct effect of the constant on M. c′ = the direct effect of X on Y. b = the direct effect of M on Y. i_Y = the direct effect of the constant on Y. [d)] Situation awareness on Mission Success mediated by Teamwork. [e)] Mission Success on Teamwork mediated by Situation awareness. [f)] Mission Success on Situation awareness mediated by Teamwork.*

Discussion

Psychometric analysis

The present study employed a version of the SMAQ but does not entirely replicate the original structure. Instead the present analysis yields a factor structure that closely replicates that of Röttger, Vetter and Kowalski (2013) who's three factors are all represented in the present study (F1, F2, and F4), and even more so O'Connor (2011) who also parcels out F3 from F2. The items of the fifth factor in the present study have elsewhere been placed under F1. However, they are very explicit in their dealing with the revealing of own short comings and invokes the language of shame, hence assigning them to a separate factor makes sense. It could also turn out that this result is due to either the sample being Norwegians, being commercial bridge officers, or both.

Hypotheses

Kirkpatrick's (1976; 2009) recommendations for training evaluation have been widely cited and has also informed the present study. Kirkpatrick (1976; 2009) conceptualizes training evaluation in four levels: 1) Satisfaction, 2) Learning, 3) Behavior change and 4) Organizational outcomes. In the present study, H1 refers to level 1) Satisfaction, H2 and H3 both concerns level 2) Learning, as Kirkpatrick collapses attitudes and knowledge into one level. However, the conceptual separation of attitudes and knowledge is a basic underpinning assumption within cognitive psychology, hence the need for measuring knowledge and attitudes separately should be self-evident. H4 and H5 are concerned with Level 3) Behavior change. Level 4 Organizational outcomes has not been measured in the present study.

H1 posited that the BRM course would be evaluated positively. H1 was fully and substantially supported. This is in accordance with considerable evidence for self-reports that participants value CRM training (DeChurch, et.al. 2010; Espevik, Saus, & Olsen, 2017).

H2 posited that the BRM course would increase the knowledge. H2 was fully and substantially supported. This finding is in accordance with some studies (Röttger et al., 2015), whereas other studies performed in the maritime domain have shown null-results and suggested that this was due to insufficient domain adaptation (O'Connor, 2011). The current support for H2 then, may be seen as supporting O'Connnor's (2011) claim about the importance of domain adapted training.

H3 posited that the BRM course would improve attitudes. H3 was partly supported. As with knowledge, the research literature gives hope, although the most relevant studies for the maritime domain has shown null results (O'Connor, 2011; Röttger, 2015). Again, the adaptation of training to the maritime domain in addition to the fact that the course employed extensive use of simulators in addition to classroom training may in part explain the deviating finding from previous maritime studies (Hobgood et al., 2010; Shapiro et al., 2004). Also, there could be sample particularities at play, since both the other maritime studies employs naval bridge officers. Concerning the varying attitudinal improvement between the five attitude factors, it has been found previously that the different attitudes measured by CMAQ show different proneness to attitude change (Helmreich et al., 1999). Hence, structural differences between the SMAQ versions can also be part of the explanation.

H4 posited that the BRM course would improve Behavior. Sadly, H4 received no support, as the present study failed to produce evidence of behavioral support. This is in accordance with previous studies (O'Connor, 2011; Röttger et al., 2016). However, there are some positive tendencies in the present study that might have proven statistically significant with a larger sample. Here it is prudent to remind the reader that at team level, the present sample is down to N=24/26 for the behavior measures. This is a mere third of the sample size used to test the other hypotheses.

In addition, there are several complicating factors contributing to the present result: 1) For the training to show behavioral change on the team level, the individuals would have to change together in such a way that the group dynamic would change. This could easily be thwarted by other team members either being very good or very poor at the onset. 2) Following requests of the ship owner, the BRM training was carried out in one standard fashion for bridge officers from two vastly different shipping trades. Hence, although random counterbalance of the pre- and post-scenarios were employed, the dramatically different technical skills involved will have diluted any effect of increased non-technical skills on actual behavior. 3) Two out of three raters were professional instructors with commercial maritime practice, without special training in human factors assessment. It is possible that this has contributed to random error, again diluting any effect. 4) Since all assessment were done from video recordings to prevent researcher bias, the assessors were vulnerable to varying quality of both video and audio. This may have centered the assessment scores, which then leads to smaller group differences and smaller effects. 5) Changes in the simulator staff, and occasional technical issues, prohibited full experimental control in the conduction of the scenarios. It should be noted that these are minor issues unproblematic to normal simulator training, but which serves to introduce potentially substantial random error in experimental measurements. Such random error reduces the internal validity of the experiment and would show up in reduced group differences and smaller, less statistically significant effects.

H5 posited that the relationship between Teamwork and Mission success is positively mediated by Situation awareness at each point in time. H5 was fully supported, adhering to a general input-process-output model where Teamwork is seen as a precursor to Mission success mediated by Situation awareness. (see e.g. Espevik et al., 2006; Espevik et al., 2011a; 2011b; Stout, Cannon-Bowers, Salas, & Milanovich, 1999). That the relationship was fully mediated also underlines the importance of including Situation awareness when assessing outcomes of BRM training, as it gives fuller understanding of the performance of a bridge crew than exclusively focusing on Teamwork and Mission Success.

Limitations

The present sample size is problematic in terms of statistical power for all analysis except the hypothesis testing of H1 regarding training satisfaction and H2 regarding knowledge improvement. Concerning the study's generalizability, the sample is also limited in representing mostly two narrow trades within commercial shipping. However, tugboats and offshore support vessels may be regarded as extremes on a continuum from small to large crews that will envelope much of the crew sizes on most commercial shipping. Also, the inclusion of both offshore maritime operations and coastal navigation in the scenarios makes the sampled exercises generalizable to much commercial shipping. In addition, substantial heterogeneity in the sample in terms of both technical- and non-technical skills have increased random error, which would suggest the sample not to be overly special.

Being part of a commercial course, the experimental control for the present study was less strict than recommended, which constitutes a threat to internal validity, increasing random measurement error and the likelihood of type-two bias in hypothesis testing. However, this lack of experimental control is typical of commercial BRM training, and as such this support claims of external validity.

According to Kirkpatrick's (1976; 2009) recommendations, the study should have included measures of organizational level change. This was dropped early in the design faze as it became clear that it was logistically impossible to follow the performance of the officers as part of their actual bridge teams since they were not coursed together as teams. Furthermore, the evident challenges for revealing any behavioral change in the controlled simulator environment precluded any reasonable expectations of evaluating the same at organizational level (Röttger et al., 2016).

Furthermore, it may be discussed whether long term behavior change belongs to Kirkpatrick's level three or level four. However, as neither the present study nor previous studies have been able to demonstrate even short term behavior change in controlled experimental settings, any long term effects - especially in natural the officers natural environment at work on their own vessels – is unlikely with present training schemes and experimental designs.

Effects of history and training are important to consider for a study like the present where there is no control group not receiving treatment. However, for the theoretical test, the questions are interspersed with four days and no discussion of correct answers were treated, so learning as a bias is unlikely. For the behavioral measure, two different scenarios were used that offered little concrete transference enabling success in the next scenario, other than actually improving resource management skills. Concerning historic effects, five days is a short period of time that is not likely to produce other effects in parallel to the experiment. And lastly, distributing the experiment over eight different weeks over six months with all experimental conditions represented each week should render small chances of random historic changes during the experiment. In fact, no major incidents were reported with shipowner during the six month period of experiment trials.

Regarding H5, a word of caution is necessary in relation to the limitations of OLS regression analyses. It cannot test the causality of the modeled structures, meaning that the directions of relationships given in the models cannot be taken for granted. Alternative causal directions could be possible. For instance, a model was tested that Mission Success had a positive relation to Teamwork, positively mediated by Situation awareness. However, all possible mediations combining these three variables were tested, and none of the alternative mediation models proved significant. A more problematic bias could be introduced by expectations from the observers who were not blind to that particular element of scoring. In other words, they could subconsciously see the different behavior scores of one session in relation. However, in order to fully experimentally control for this, separate observers would have been necessary for the different behaviors.

Regarding the instrument validity, the SMAQ questionnaire proved problematic: The factor loading could still be clearer, with more balance in the number of items for each factor and with higher Cronbach alphas. This may have contributed to random error. Seeing that the two most obviously comparable studies uses two different versions which also have some structural issues, a revision is called for (O'Connor, 2011; Röttger, et al., 2013).

General discussion

Limitations notwithstanding, the present study provides supporting evidence that it is indeed possible to achieve results in resource management training within the maritime domain, in the way it has been shown in other domains (Hughes et al., 2016; Weaver et al., 2010; Espevik, Saus & Olsen, 2017). When the evidence from the literature within the maritime domain has been meager so far, it should perhaps be viewed in connection with certain attributes of the maritime domain: 1) The acceptance of risk in shipping is higher than in other domains such as aviation and medicine (IMO, 2011). 2) The level of academic training amongst practitioners and instructors are comparably low. 3) The adaptation of BRM training within the commercial maritime sector is typically viewed to be relevant for any bridge officer situated within commercial shipping, be it at a one-man fishing vessel, a large super tanker, other tankers- or offshore-related vessels doing highly specialized operations, or different kinds of rigs. The medical equivalent would be to say that the resource management training could use the same scenarios for emergency room surgeons, brain surgeons, and heart surgeons alike. 4) The requirements for resource management training in the maritime domain remain far less strict than in other domains. Currently, the typical course requirements following the 2010 Manilla amendments to the STCW regulations where BRM was for the first time made compulsory, is for a generic three- or four-day BRM course (International Maritime Organization, 2011).

The fact that the present study is capable of showing some improvement in both knowledge and attitude, as well as promising tendencies regarding behavior as well as evidence for the behavioral connection between Teamwork, Situation awareness, and Mission success, supports previous calls for domain-relevant training and the use of simulator training in addition to classroom training (Helmreich et al., 1999; Håvold et al.., 2015; O'Connor, 2011; Röttger et al., 2013; 2016).

It has been noted previously that the level of research on themes relevant for resource management training has been low within the maritime domain, theoretically reducing the value of BRM training (Hetherington, Flin, & Mearns, 2006). Lastly, previous calls to the importance of scientific evaluation of training programs within resource management are as valid as ever, also including the maritime domain (Flin et al., 2008). Furthermore, the principles laid out for such evaluations are readily available and relevant for evidence-based resource management training for commercial bridge officers (Kirkpatrick, 1976; 2009; Espevik et.al. 2017). However, it is presently unclear how organizational change - the last level of evidence in prescribed by Kirkpatrick's classic model – should be measured in practical way as an experimental outcome.

Conclusion

The present study produces evidence that adapted resource management training employing full-mission simulators can indeed be performed to the satisfaction of participants and improve relevant knowledge and attitudes within the maritime domain. Also, there are promising trends, although not statistically significant, that it is possible to improve relevant behavior as measured in a controlled simulator environment. Finally, the present study produces correlational support for the association between Teamwork and Mission success, positively mediated by Situation awareness. Researchers and practitioners alike should aim to specialize the BRM-training for more focused maritime trades and operations. Furthermore, the developments are needed for measuring non-technical skills and safety-relevant behavior at the organizational level in order to connect this outcome to the evaluation of training regimes.

Key Points

- Safety relevant knowledge and attitudes can be improved by Bridge Resource Management (BRM) training.
- Training must be adapted to maritime domain to be effective
- Behavioral change was not significant but could be probably achieved with better customized training design.
- Established resource management training theory concerning mechanisms of team behavior was supported.
- Teamwork was related to Mission Success, mediated by Situation awareness.

Acknowledgments

The authors wish to thank Simsea simulation center and the ship owner for facilitating the present study. The Norwegian Research Council and nine companies in the offshore business in Norway have funded the RISKOP project; Managing Risk in Offshore Operations (grant no.: 225311/O70) at Stord/Haugesund University College, now University of Applied Sciences of Western Norway. UH-Nett Vest, an organization of cooperating Western Norwegian Universities and Colleges, also contributed to the funding of the article. We acknowledge the work of Professor Jan R. Jonassen of the Western Norway University of Applied Sciences (HVL) for finance and facilitation of this research, and Technician Morten Mæland from HVL for facilitation of the exercise recordings.

Statement regarding conflict of interest:

The Main author is employed at the simulation center where the evaluated training has taken place. However, he has not been involved as assessor, and the objective treatment of data has been monitored by the other authors who have no financial ties with the training provider or the ship owner.

References

Andersen, H.B., Garay, G. and Itoh, K. (1999). Survey data on mariners: Attitudes to safety Issues. Technical Report I-1388, Systems Analysis Department, Risø National Laboratory, DK-4000 Roskilde, Denmark.

Baron, R. B., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Converse, S., Cannon-Bowers, & J. A., Salas, E. (1993). Shared mental models. In J. N Castellan (Ed.), *Individual and group decision making* (pp. 221-246). Hillsdale, NJ: Erlbaum.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology, 95*(1), 32.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Harcourt Brace Jovanovich College Publishers.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems.*Human Factors, 37* (1), 32-64.

Endsley, M. R. (2015). Situation awareness: operationally necessary and scientifically grounded. *Cognition Technology & Work, 17*, 163-167. doi: 10.1007/s10111-015-0323-5.

Eraut M. (2000). Non-formal learning and tacit knowledge in professional work. *Br J Educ Psychol.*;70:113-136.

Espevik, R., Johnsen, B.H., Eid, J., & Thayer, J. (2006). Shared Mental Models and Operational Effectiveness; Effects on performance and team processes in a submarine attack team. *Military Psychology, 18*, 23-36

Espevik, R.E., Johnsen, B.H., &Eid, J. (2011 a). Communication and Performance in Co-Located and Distributed Teams: An Issue of Shared Mental Models of Team Members? *Military Psychology, 23*, 616–638

Espevik, R.E., Johnsen, B.H., &Eid, J. (2011 b). Outcomes of shared mental models of team members in cross training and high intensity simulations" *Journal of cognitive engeneering and decision making, 5*, 352–377.

Espevik, R., & Olsen, O. K. (2013). A new model for understanding teamwork onboard: The shipmate model. *International maritime health, 64*(2), 89-94.

Espevik, R., Saus, E. R., & Olsen, O. K. (2017). Exploring the core of crew resource management course: speak up or stay silent. *International maritime health, 68*(2), 126-132.

Flin, R. H., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: a guide to non-technical skills*. Ashgate Publishing, Ltd.

Hansen, H. L. (1996). Surveillance of deaths on board Danish merchant ships, 1986-93: implications for prevention. *Occupational and environmental medicine, 53*(4), 269-275.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York, NY: The Guilford Press.

Helmreich, R.L. and Merritt, A.C. (1998). Culture at work in aviation and medicine: National, organizational and professional influences. Ashgate, Aldershot, UK.

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *The international journal of aviation psychology, 9*(1), 19-32.

Hetherington C, Flin R, Mearns K (2006) Safety in shipping: the human element. J Safety Res 37(4): 401–411. https://doi.org/10.1016/j.jsr.2006.04.007.

Hobgood, C., Sherwood, G., Frush, K., Hollar, D., Maynard, L., Foster, B., ... & Taekman, J. (2010). Teamwork training with nursing and medical students: does the method matter? Results of an interinstitutional, interdisciplinary collaboration. *Quality and Safety in Health Care*, qshc-2007.

Hughes, A. M., Gregory, M. E., Joseph, D. L., Sonesh, S. C., Marlow, S. L., Lacerenza, C. N., ... & Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare.

Håvold, J. I., Nistad, S., Skiri, A., & Ødegård, A. (2015). The human factor and simulator training for offshore anchor handling operators. *Safety Science, 75*, 136-145.

International Maritime Organization (2011). STCW Convention and STCW Code. International Convention on Standards of Training, Certification and Watchkeeping for seafarers. Including the 2010 Manilla Amendments. London: International Maritime Organization International Maritime Organization.

Wick, K. E., (2002). Safer shipping demands safety culture. Paper presented at the World Maritime Day.

Kelley, K. (2018). *The MBESS R Package.* Date/ Publication 2018-01-10 23:37:02 UTC.

Kirkpatrick DL. (1976). Evaluation. *in* CraigRL (Ed.), *Training and development handbook.* (*pp.* 301–319). *New York: McGraw-Hill*

Kirkpatrick, D. L. (2009). *Implementing the Four Levels: A Practical Guide for Effective Evaluation of*

Training Programs: Easyread Super Large 24pt Edition. ReadHowYouWant. com.
Lewin, K. (1943). Forces behind food habits and methods of change. *Bulletin of the national Research Council, 108*, 35-65.
Lewin, K. (1947). Group decision and social change. *Readings in social psychology, 3*, 197-211.
Mjelde, F. V., Smith, K., Lunde, P., & Espevik, R. (2016). Military teams–A demand for resilience. *Work*, (Preprint), 1-13.
Musson, D. M., & Helmreich, R. L. (2004). Team training and resource management in health care: current issues and future directions. *Harvard Health Policy Review, 5*(1), 25-35.
O'Connor, P. (2011). Assessing the effectiveness of bridge resource management training. *International Journal of Aviation Psychology, 21*, 357–374.
O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. A. (2008). Crew resource management training effectiveness: A meta-analysis and some critical needs. *The International Journal of Aviation Psychology, 18*(4), 353-368.
Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers, 36*, 717–731.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.
Rouse, W. B. and Morris, N. M. (1986). On looking into the black box: Prospects and limits in search for mental models. *Psychological Bulletin, 100*, 349-363.
Rutherford, J. S., Flin, R., & Mitchell, L. (2012). Non-technical skills of anaesthetic assistants in the perioperative period: a literature review. *British journal of anaesthesia, 109*(1), 27-31.
Röttger, S., Vetter, S. and Kowalski, J.T., (2013). Ship Management Attitudes and Their Relation to Behavior and Performance, Human Factor, 55(3), pp. 659-671.
Röttger, S., Vetter, S., & Kowalski, J. T. (2016). Effects of a classroom-based bridge resource management training on knowledge, attitudes, behaviour and performance of junior naval officers. *WMU Journal of Maritime Affairs, 15*(1), 143-162.
Salas, E., Cannon-Bowers, J., Rhodenizer, L., & Bowers, C. (1999). Training in organizations: Myths, misconceptions, and mistaken assumptions. Research in Personnel and Human Resource Management, 17, 123–161.
Salas, E., DiazGranados, D., Weaver, S. J., & King, H. (2008). Does team training work? Principles for health care. *Academic Emergency Medicine, 15*(11), 1002-1009.
Salas, E., Sims, D. E., & Burke, C. S., (2005). Is there a "big five" in teamwork. *Small Group Research, 36*, 555-599.
Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. Psychological science in the public interest, 13(2), 74-101.
Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4 i environments. *Applied Ergonomics, 37*, 225-238.
Shapiro, M. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., ... & Jay, G. D. (2004). Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum?. *Quality and Safety in Health Care, 13*(6), 417-421.
Sneddon, A., Mearns, K., & Flin, R. (2006). Situation awareness and safety inoffshore drill crews. *Cognition, Technology and Work, 8,* 255-267.
Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). Washington, DC: American Sociological Association.
Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(1), 61-71.
Weaver, S. J., Lyons, R., DiazGranados, D., Rosen, M. A., Salas, E., Oglesby, J., ... & King, H. B. (2010). The anatomy of health care team training and the state of practice: a critical review. *Academic Medicine, 85*(11), 1746-1760.
Weick, K. E. (2002). The reduction of medical errors through mindful interdependence. In M. M. Rosenthal & K. M. Sutchcliffe (Eds), Medical error: What do we know? What do we do? (pp.177-199). San Francisco: Jossey-Bass.