

# Effects of a classroom-based bridge resource management training on knowledge, attitudes, behaviour and performance of junior naval officers

Stefan Röttger · Saskia Vetter · Jens T. Kowalski

Received: 7 July 2014 / Accepted: 11 December 2014 / Published online: 27 January 2015  
© World Maritime University 2015

**Abstract** This study assesses the effectiveness of classroom-based bridge resource management (BRM) training for junior naval officers, in which general principles of human behaviour and performance in teams and under stress were conveyed. Although BRM training is recommended by the International Maritime Organization and is increasingly common in seafaring, very little is known about whether the adaptation of crew resource management (CRM) training from aviation to the maritime domain has been successful and what type of training is effective. A study with a quasi-experimental, two-factorial mixed design was conducted with BRM training as the between factor and time as the within factor. For 117 study participants, evaluation criteria were assessed on all levels as defined by Kirkpatrick (Train Dev J, 178–192 1979): subjective training evaluation, knowledge, attitudes and behaviour as well as performance while commanding a vessel during a real-world exercise. BRM participants showed better subjective training evaluations and more BRM-related knowledge than controls. Training did not produce differences between groups regarding BRM-related attitudes, the demonstration of non-technical skills or the overall success in the real-world exercise. Overall, BRM training

---

S. Röttger (✉)  
German Armed Forces Office, Military Psychology Research Unit,  
c/o Helmut-Schmidt-University  
Holstenhofweg 85, 22403 Hamburg, Germany  
e-mail: stefanroettger@bundeswehr.org

S. Vetter  
Department of Naval Psychology and Ergonomics, German Naval Medical Institute,  
Kopperpähler Allee 120, 24119 Kronshagen, Germany

J. T. Kowalski  
German Armed Forces Office, Department of Military Psychology  
Fontainengraben 150, 53123 Bonn, Germany

effectiveness was rather low, which can most probably be attributed to the focus of training on generalizable knowledge, skills and attitudes at the expense of their specific application to the context of the real-world exercise. In the design of BRM and CRM training courses alike, the effective application of general principles to a given context must be defined, and the application must be emphasised during training delivery.

**Keywords** Team training · Transfer of training · Maritime systems · Nautical teams · Training evaluation

## 1 Introduction

With the increasing number of work teams controlling complex technical systems in modern industries, enhancing team performance by means of team training became a major concern in many organizations (Salas et al. 2008). One kind of training that emerged with the increasing need for team training in high technology domains is crew resource management (CRM) training. The purpose of CRM training is to impart knowledge and skills that are, in addition to technical expertise, necessary to accomplish the tasks of a team safely and successfully. The focus of CRM training varies, but usually involves topics such as communication, coordination, cooperation and mutual support, leadership, decision making, situation awareness, limitations of human performance, stress and fatigue.

Crew resource management training originated in aviation (Helmreich et al. 1999) and is now mandatory for commercial pilots in many countries, including the European Union (European Commission 2012) and the United States of America (Federal Aviation Administration 2013). The idea and concepts of CRM training have been transferred and adapted to other areas of teamwork in high-risk domains, such as seafaring, nuclear power production, offshore oil and gas production, the military and, above all, healthcare (Flin et al. 2003).

Reviews and meta-analyses show that CRM training is generally effective (Salas et al. 2008), it is positively perceived by participants and consistently improves teamwork-related attitudes (Salas et al. 2006b; O'Connor et al. 2008). While CRM training induces changes in knowledge and behaviour, effect sizes vary considerably (Ibid.) and thus seem to depend more strongly on the focus and design of the particular training course.

The focus of the present article is CRM effectiveness in seafaring teams. In a recent analysis of 27 collision accidents that occurred in the years between 1998 and 2012, Chauvin et al. (2013) reported crew resource management deficits in 38 % of the involved bridge teams. Just like in aviation, mishaps and accidents at sea led to the recommendation of non-technical skills training, which started to develop independently of aviation in merchant shipping in the late 1970's (Barnett et al. 2003). From the 1990s, training was strongly influenced by CRM training common in aviation, and the term Bridge Resource Management (BRM) training was coined for CRM training of seafaring teams. From a regulatory perspective, bridge resource management competencies in general, and leadership and teamwork in particular, have been

stressed in the Standards of Training, Certification and Watchkeeping for Seafarers (STCW) Code since 2010 (IMO 2010). Ship's masters, officers and engineers are required by this international convention to achieve a sufficient level of proficiency in these competencies. For certification purposes, seafarers can demonstrate that they fulfil these requirements by undergoing bridge resource management training, but confirmed in-service experience can serve as evidence of sufficient bridge resource management skills as well. As bridge resource management training is not mandatory, this kind of training has been evolving slower than CRM, is not as common place as CRM and less research has been conducted on training effectiveness than in aviation. Although there are some papers that report BRM training effects (Fonne and Fredriksen 1995; Byrdorf 1998; Brun et al. 2005; O'Connor 2011), there are just as many papers calling for more and better research into BRM training (Barnett et al. 2003; Hetherington et al. 2006; Salas et al. 2006b; O'Connor et al. 2008).

The earliest report of BRM training effects can be found in a conference paper by Fonne and Fredriksen (1995), who describe the introduction of BRM training for navigational officers of high-speed vessels in Norway. The authors report questionnaire results indicating that compared to pre-seminar assessments, navigators preferred less autocratic leadership styles, would reveal personal stress more readily to crew members and would abide by higher safety margins after the seminar. However, the actual size and significance of the reported effects remain rather elusive because few descriptive statistics and no inferential statistics are given. Moreover, the simple pre-post design without a control group makes it impossible to infer any causal relationship between the described changes in attitude and the participation in the BRM seminar.

Byrdorf (1998) reports the effects of a BRM training programme that was introduced by the MAERSK shipping company in 1994 in order to improve leadership, assertiveness, communication, team work and stress coping among the nautical and engineering personnel of their vessels. BRM training consisted of a classroom course lasting 4 days and 3 days of simulator exercises with debriefings. Regarding the effectiveness of this training, Byrdorf cites safety and damage records of MAERSK from 1992 and 1996 which show a marked reduction of averages, lost-time accidents and insurance premiums. Unfortunately, owing to the missing control group, it is again unclear whether the positive trends in safety records can be attributed to the BRM training programme, or whether these trends were the result of other factors, e.g., technical and organizational measures that may have been implemented during this time period.

An experimental control-group design was implemented by Brun et al. (2005) to investigate the effectiveness of a BRM training course that lasted 1 week and was composed of lectures, case studies, practical planning tasks and simulator rides. Differences in shared mental models and several aspects of performance between naval bridge teams with and without BRM training were analysed. The results are inconclusive, because the number of teams (two in each condition) was not sufficient to statistically verify any training effect. On a descriptive level, the changes observed after the BRM training course were very small and not systematic in any direction.

A more extensive sample was used by O'Connor (2011) in a survey study on BRM-related knowledge and attitudes of 166 surface warfare officers of the U.S. Navy (parts of the data were previously reported in Carter-Trahan (2009)). Knowledge and attitudes were compared between participants that had already attended a BRM course (86 %) and those who had not at the time of the study. No significant differences were found. The ineffectiveness of the BRM training course under study, which comprised of 14 h of classroom instruction and 20 h of simulator training, was explained by the fact that training was not based on a training needs analysis and did not cover central issues of bridge resource management such as decision making, situation awareness, stress and fatigue.

All in all, the results regarding BRM effectiveness are inconclusive: Early studies report positive changes in individual attitudes and mishap rates after the introduction of BRM training, but due to the lack of control groups in these studies, it is not clear whether the observed changes were actually caused by the implementation of BRM courses. The only study permitting statistical analyses of differences between trained and untrained study participants suggests that the BRM training course in question did not effectively influence BRM-related knowledge and attitudes. In this latter study, there was no experimental manipulation of BRM training participation, so equivalence of BRM and non-BRM groups could not be ensured.

The inconclusiveness of the empirical findings on BRM training effectiveness is owed to a number of gaps in the research literature: So far, there are no reports on BRM training effects on behaviour and performance of seafarers, and there is no study combining an experimental control group and a sufficiently large sample size to justify sound conclusions on training effectiveness. The purpose of the study presented here was to fill these gaps in the research literature and to provide a firm data base for future meta-analysis by employing an experimental control group design, assessing data regarding reactions, knowledge, attitudes, behaviour and performance of study participants, and by ensuring the equivalence of experimental groups in a sufficiently large sample.

This methodological approach was used to evaluate a 5-day, classroom-based BRM course which was on trial at the German Naval Academy as part of the leadership studies of junior naval officers. As real-world exercises in navigating and commanding vessels are conducted during the leadership studies at the German Naval Academy, this offered an opportunity to assess behaviour and performance of study participants under realistic conditions and to determine whether the contents of the classroom-based training are transferred and applied to the actual operational setting by the participants. The exact details of the BRM training under study, of the research methods and of the study design, are described in the next section.

## 2 Methods

### 2.1 Sample

One hundred seventeen junior naval officers participated in this study during their leadership studies at the German Naval Academy. Of these 117 participants, 57

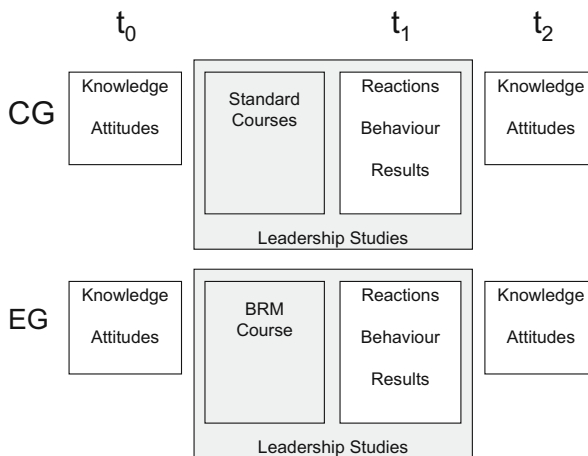
belonged to the experimental group and received BRM classroom training. The average age was 24.7 years ( $sd = 2.3$ ), 21 (18 %) study participants were female. The distribution of military ranks was 91 (78 %) Lieutenants (NATO rank OF-1), 15 (13 %) Officer Cadets (NATO rank student officer) and 11 (9 %) Officer Designates (NATO rank OF-D). Some participants were not available at all data acquisition stages, so sample sizes reported in subsequent analyses and results may be lower than the total number of participants.

## 2.2 Design and procedure

Figure 1 gives an overview of the design and the stages of data acquisition. The study follows a two-factorial mixed design with BRM training as the between factor (course without or with BRM contents) and time as a within factor (pre- or post-course). For some variables, only post-training data could be assessed due to limited resources (behaviour and performance in real-world exercise) or because a pre-test would not yield any sensible result (reactions to training).

Participants were assigned course wise to conditions. Out of six consecutive courses in leadership studies, three courses (1, 3 and 6) were assigned to the experimental group and received BRM training. Assignment of the individual courses to experimental group or control group was determined by the German Naval Academy on the basis of organisational constraints regarding the availability of resources to accommodate the BRM training in the schedule of a particular course. These constraints were not related to personal or demographic characteristics of the participants, so group assignment can be considered to be pseudorandom.

During week four of the leadership studies, the experimental group participated in a 5-day BRM classroom training course that was designed and conducted by officers on the basis of their experience in implementing a CRM program for helicopter pilots of the German Armed Forces. The training objective was to impart knowledge,



**Fig. 1** Experimental design. *CG* Control group. *EG* Experimental group

improve skills and change attitudes regarding decision making, leadership, communication, coordination, performance under stress and situation awareness (see Table 1 for example contents). Training contents were not tailored to a certain task, but general principles of human behaviour and performance were presented and participants were asked to draw conclusions for their own duties. Instructional methods in this training course are comprised of lectures, (video) presentations of example incidents and accidents from aviation and seafaring, group discussions and classroom exercises in fictitious scenarios.

Participants in the control group completed the standard schedule of course week four. Standard courses are comprised of topics such as military and legal foundations

**Table 1** Overview over BRM course contents

Topic	Example Contents
BRM foundations	Origin and development of CRM/BRM trainings contribution of human behaviour to safety definition of safety-relevant non-technical skills
Decision making	Types of decisions, e.g. routine versus nonroutine cognitive processes and common errors in decision making FORDEC: a prescriptive model for nonroutine decisions
Leadership	Leadership styles, e.g. task oriented versus employee oriented consequences of different authority gradients for safety assertiveness
Communication	Non-verbal aspects of communication impact of stress on verbal expressiveness effective feedback
Cooperation	Basic cooperative patterns in social interactions cooperation as a determinant of team effectiveness social conflicts and methods for conflict resolution
Attitudes and motivation	Basic models of human motivation (e.g. expectancy-value theory) safety consequences of implicit motives (e.g. need for achievement) safety culture and safety-critical attitudes in seafaring
Performance under stress	Typical stressors and stress effects on human performance sleep-related and task-related fatigue classification of human errors and related countermeasures
Situation awareness	Models of situation awareness and related cognitive processes indicators of impaired situation awareness measures to improve situation awareness

of leadership, military studies, sports, stress, communication and the preparation and presentation of military situation reports.

At the beginning of the leadership studies ( $t_0$ ), participants were informed about the purpose and procedure of the research project and written informed consent was obtained. Questionnaire data were acquired at three stages. At stage  $t_0$ , a pre-test of BRM-related knowledge and attitudes was conducted. At stage  $t_1$  (after week four of leadership studies), reactions regarding the courses of the previous week (BRM or standard) were collected. In addition, behaviour and performance during a real-world exercise were observed and rated. During this exercise, each participant was required to command and navigate a boat with several crew members on the Flensburg Firth and to accomplish a task in a fictitious humanitarian aid scenario in a politically unstable region (see Fig. 2). For example, routes for safe navigation or places suitable to safely unload goods had to be found and secured. The exercise was a practical part of leadership studies, so instructors emphasised the importance of leadership and teamwork in briefings and debriefings. Data acquisition stage  $t_2$  followed 3 weeks after  $t_1$  at the end of the leadership studies. At this stage, a post-test of BRM-related knowledge and attitudes among study participants was carried out.

### 2.3 Measures

Selection of dependent variables was based on Kirkpatrick's classification of evaluation criteria (Kirkpatrick 1979) and on an augmented framework of this classification as proposed by Alliger et al. (1997). Dependent variables were thus assessed on each of the four levels of training success: *reactions*, i.e. the participants' subjective evaluation regarding the utility and the appeal of training; *learning*, which comprises improved knowledge and changes in attitude that may be caused by training; *behaviour* of the study participants on the job, and results of their behaviour, which will be referred to as *performance* in this article.



**Fig. 2** Study participant commanding his boat in the real-world exercise

### 2.3.1 Reactions

The questionnaire for the assessment of trainee reactions was based on the results of Staufenbiel (2000) and Holgado Tello et al. (2006), who independently report that trainee reactions usually comprise three areas of training evaluation. The first area regards the quality of organisation and presentation of training content. The second area regards the degree of interest and the relevance of the training content for the trainee's personal needs and field of work. This corresponds to Kirkpatrick's utility reaction. The third area is a global assessment of training quality, which in some points resembles the affective reaction in accordance with Kirkpatrick.

The questionnaire in this study is a modification of the evaluation form for university courses published by (Staufenbiel 2000), which comprises these three aspects of trainee reactions. In some items, the wording was slightly adapted to the terminology in use at the German Naval Academy. Questions regarding the behaviour of individual trainers or the quality and quantity of resources (e.g. handout copies) were excluded because they were not within the focus of this study. Three items were added to permit a direct assessment of the overall evaluation, the affective reaction and the utility reaction with regard to training. Scales and items contained in the resulting questionnaire are listed in Table 2.

Agreement or disagreement with the statements is assessed with a 5-point rating scale ranging from 1 (completely disagree) to 5 (completely agree). Mean values

**Table 2** Scales and items of the reactions questionnaire

Scale	Item
Organisation and presentation	The course content follows a clear structure. The course gives a good overview of the subject. Too little explanatory or supplementary information was given on the course topics. (-)
Interest and relevance	The course content and methods were interesting. The course content is useful for my duties. The applicability and utility of the course content were not sufficiently made clear. (-) I can apply the acquired knowledge and skills in my (prospective) duties.*
Global evaluation	I liked the course.* I acquired theoretical knowledge during the course. I acquired practical skills during the course.* What mark would you give the course? (-)

Notes: English translation of German items and scale labels. Items not marked with an *asterisk* are from a course evaluation form in Staufenbiel (2000). Items marked with an *asterisk* were added. *Minus sign* indicates reversed polarity.

greater than three express a rather positive, values smaller than three a rather negative evaluation. Ratings of items with reversed polarity are reversed before further data processing.

### 2.3.2 Knowledge

BRM-related knowledge was assessed with 13 open questions on decision making, leadership, communication, cooperation, performance under stress and situation awareness. The questions were derived from the goals and content of BRM training. About half of the questions were designed to elicit declarative knowledge (e.g. “Please name and describe three non-technical skills”), while the other half aimed at procedural knowledge (e.g. “How can you make a decision in a new and complex situation?”). Answers were scored in a standardized fashion using a scoring reference. A maximum score of 50 could be obtained in the knowledge test.

### 2.3.3 Attitudes

BRM-related attitudes were assessed using the Ship Management Attitudes Questionnaire – German Navy (SMAQ–GN, Röttger et al. (2013)), which is based on the Cockpit Management Attitudes Questionnaire (CMAQ, Helmreich (1984), Gregorich et al. (1990)) and its German translation by Hörmann and Maschke (1991). Internal consistencies of the SMAQ–GN scales are comparable to those of the original CMAQ and test-retest-reliabilities of the scales are between 0.74 and 0.81 (Röttger et al. 2013).

The questionnaire contains 17 items with a rating scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). Participants are asked to indicate on this rating scale whether they agree or disagree with the individual statements. Scales of the SMAQ–GN are the same as in the CMAQ, i.e. *Communication and Coordination* (COCO, 10 items), *Command Responsibility* (COMMAND, 3 items) and *Recognition of Stressor Effects* (RSE, 4 items). Scale values are calculated as the mean of all ratings of a scale. Values of items with reversed polarity (scales COMMAND and RSE) are reversed to fit the scale on which higher values are associated with more effective attitudes. Further details on the questionnaire can be found in Röttger et al. (2013).

### 2.3.4 Behaviour

The behaviour of participants commanding a boat in the real-world exercise was observed and rated by a team of three Psychologists and three senior officers of the German Navy using the NON-TECHNical Skills observation and rating system NOTECHS (van Avermaete and Kruijsen 1998; O’Connor et al. 2002). NOTECHS comprises the categories leadership, cooperation, decision making and situation awareness. Each category contains three to four elements that contribute to the overall rating of a category (e.g. anticipation in the category situation awareness). For each element, example behaviours are given to direct the observation (e.g. “Discusses contingency strategies” for anticipation). In order to increase reliability

and validity of the observational data, all observers received a comprehensive briefing on the regime of the data acquisition during the exercises, on the content, the structure and the use of the rating system NOTECHS, and on example behaviours that may occur during the exercises. In addition, the senior officers received a psychological training that acquainted them with common errors in social perception and attribution and with techniques to prevent these errors. The application of these techniques was practised in observations and evaluations of leadership and teamwork behaviours in video samples of emergency management scenarios in process control and healthcare. Observers were not informed whether the participants had undergone the standard course or BRM classroom training. During the exercise, there was one observer aboard each boat, who took notes on observed behaviours of the participant in command. Each element was rated on the basis of these observations as soon as a participant had finished his or her exercise. No rating was assigned if too few had been observed regarding the element in question. Ratings on the category level were derived from the arithmetic mean of the element-wise ratings. Ratings are scaled from 1: “Behaviour directly endangered safety or task accomplishment” over 3: “Behaviour can be improved but did not endanger safety and task accomplishment” to 5: “Behaviour optimally supports safety and task accomplishment and could be an example for others”.

### 2.3.5 Performance

The performance of participants in the real-world exercise served as a result-level evaluation criterion. Participants’ performance was determined by rating the degree of task accomplishment at the end of the exercise. A 5-point rating scale was used with 1 indicating complete failure to meet any of the task requirements, 3 indicating that a task was halfway accomplished and 5 indicating complete task accomplishment.

## 2.4 Analysis

Statistical analyses were carried out with the open source software *R* 2.11 (R Development Core Team 2010). First, the equivalence of control group and experimental group in terms of BRM-related knowledge and attitudes at  $t_0$  was tested. In case of significant differences, pre-test values would serve as a covariate in the subsequent multivariate analysis of variance, which was calculated to determine the overall effect of the factor BRM training on the 12 dependent variables (three reaction scales, four NOTECHS categories, performance, knowledge, three SMAQ–GN scales). In case of a significant multivariate effect, the significance of differences between groups in individual evaluation criteria was tested with Welch’s *t*-test. Degrees of freedom were corrected in case of unequal variances. Because pre- and post-training data were obtained for BRM-related knowledge and attitudes, within-subjects comparisons were conducted for these evaluation criteria. For significant group differences, the effect size *d* is provided. The results of the statistical analysis will be reported in the following Section 3, *Results*, whereas explanations and interpretations of these results will be provided in Section 4, *Discussion*.

### 3 Results

#### 3.1 Sample equivalence

Columns two to five of Table 3 contain means and standard deviations of the pre-test data obtained in the control group and the experimental group. Columns six to eight show the results of the statistical tests of differences between groups. Control group and experimental group did not differ significantly in their BRM-related knowledge and attitudes at the beginning of the study.

#### 3.2 Overall effect of BRM classroom training

Because 38 of the 117 cases in this sample contained missing values, the MANOVA was calculated with the remaining 79 complete cases. A significant effect of BRM training on the dependent variables was found,  $F(12, 66) = 2.04$ ,  $p = 0.034$ . Therefore, further analyses of each evaluation criterion were conducted.

#### 3.3 Reactions

Figure 3 shows a box plot of participant reactions to the BRM course and the standard course at  $t_1$ , directly after the completion of the course week. The horizontal bars indicate the median of the distribution. Boxes cover the central 50 % of the data range and vertical lines cover observed values of up to 1.5 times the central data range. Individual values beyond that point are represented by dots.

Most evaluations were rather positive ( $> 3$ ). The mean global evaluation of the standard course was 3.36 ( $sd = 0.66$ ), the global evaluations of the BRM course averaged to 3.59 ( $sd = 0.64$ ). Mean relevance ratings were 3.39 ( $sd = 0.69$ ) for the standard course and 3.74 ( $sd = 0.66$ ) for the BRM course. Organization and presentation of course contents received an average rating of 3.64 ( $sd = 0.57$ ) in the standard course and 3.93 ( $sd = 0.63$ ) in the BRM course.

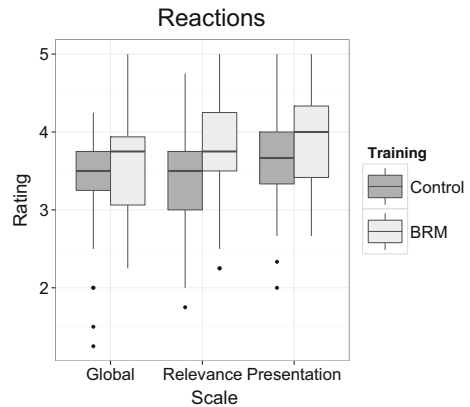
Significant differences between evaluations of BRM and standard courses were found for scales interest and relevance,  $t(102) = -2.6$ ,  $p = 0.011$ ,  $d = 0.51$  and

**Table 3** Differences between control group (CG) and experimental group (EG) in BRM-related knowledge and attitudes at  $t_0$

	CG		EG		$t$	df	$p$
	$\bar{x}$	$sd$	$\bar{x}$	$sd$			
Knowledge	13.8	5.1	14.2	4.3	-0.44	107	0.659
COCO	4.1	0.5	4.1	0.3	0.92	95.2	0.361
COMMAND	3.4	0.7	3.4	0.6	-0.15	107	0.879
RSE	3.2	0.7	3.2	0.6	-0.36	107	0.720

Notes: Knowledge: number of correct facts in knowledge test. COCO, COMMAND, RSE: scale values of SMAQ-GN.

**Fig. 3** Mean rating of standard course (dark grey) and BRM course (light grey) by study participants

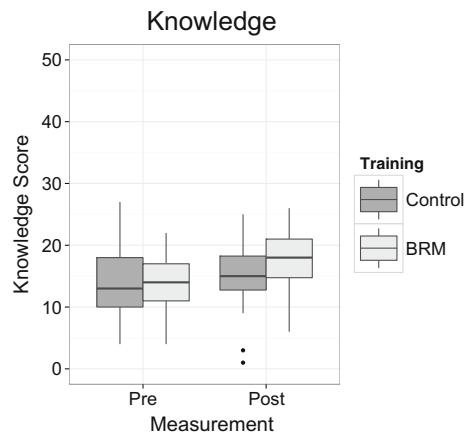


organisation and presentation,  $t(102) = -2.39$ ,  $p = 0.019$ ,  $d = 0.47$ . Global evaluations of BRM and standard courses did not differ significantly,  $t(102) = -1.79$ ,  $p = 0.077$ .

### 3.4 Knowledge

Figure 4 shows that participants of the control group achieved nearly the same scores in the knowledge test before (mean 13.8,  $sd = 5.1$ ) and after (mean 13.6,  $sd = 6.4$ ) their leadership studies. Accordingly, the  $t$  test comparing these means was not significant,  $t(42) = -0.8$ ,  $p = 0.437$ . In the experimental group, mean knowledge scores increased from 14.2 ( $sd = 4.3$ ) to 17.4 ( $sd = 4.8$ ) over the course of the leadership studies. This difference was statistically significant,  $t(48) = -5.7$ ,  $p < 0.001$ . As a result of this increase, the experimental group had significantly more BRM-related knowledge than the control group after the leadership studies,  $t(92) = -2.8$ ,  $p = 0.005$ ,  $d = 0.58$ , while there was no difference between groups when leadership studies commenced (see Table 3). With a mean score of 17.4, participants of the

**Fig. 4** BRM-related knowledge in the control group (dark grey) and the experimental group (light grey) before and after leadership studies



experimental group achieved on average 34.8 % of the maximum score of 50 in the knowledge test.

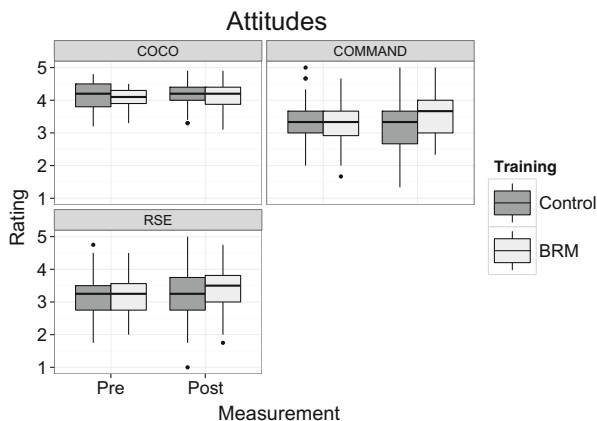
### 3.5 Attitudes

Figure 5 gives an overview of the scale values from the SMAQ–GN in both groups before and after the leadership studies. Means and standard deviations of the scales at  $t_0$  are listed for both groups in Table 3. At  $t_2$ , at the end of the leadership studies, average scale values for communication and coordination (COCO) was 4.2 ( $sd = 0.4$ ) in the control group and 4.1 ( $sd = 0.4$ ) in the experimental group. Averages of the scale COMMAND were 3.3 ( $sd = 0.8$ ) in the control group and 3.5 ( $sd = 0.6$ ) in the experimental group. Mean values of the items regarding the recognition of stressor effects (scale RSE) were 3.2 ( $sd = 0.8$ ) in participants of the control group and 3.4 ( $sd = 0.7$ ) in participants of the experimental group.

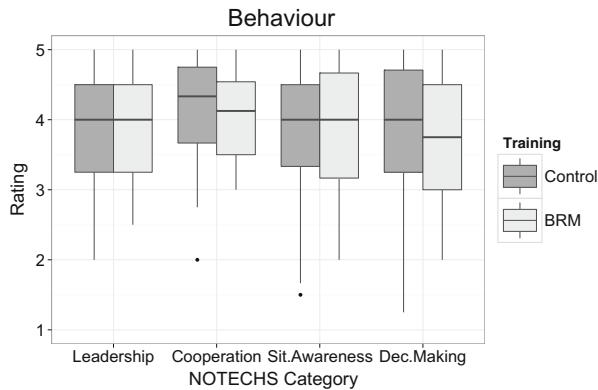
Statistical tests of the differences between groups at  $t_2$  did not yield significant results. Test statistics were  $t(99) = 0.7$ ,  $p = 0.489$  for scale COCO,  $t(99) = -1.4$ ,  $p = 0.17$  for scale COMMAND and  $t(99) = -1.0$ ,  $p = 0.317$  for scale RSE. Moreover, neither control group nor experimental group showed any significant change of attitudes over the course of the study. In the control group, results of comparisons between scale values at  $t_0$  and  $t_2$  were  $t(47) = -0.2$ ,  $p = 0.804$  for scale COCO,  $t(47) = 0.8$ ,  $p = 0.427$  for scale COMMAND and  $t(47) = 0.3$ ,  $p = 0.792$  for scale RSE. In the experimental group,  $t$  tests for dependent samples revealed values of  $t(47) = -1.5$ ,  $p = 0.152$  for COCO,  $t(47) = -1.3$ ,  $p = 0.195$  for COMMAND and  $t(47) = -1.9$ ,  $p = 0.062$  for RSE.

### 3.6 Behaviour

Figure 6 depicts the NOTECHS ratings collected in the control group and the experimental group during the real-world exercise in commanding a boat. Leadership skills



**Fig. 5** Ship management attitudes in the control group (dark grey) and the experimental group (light grey) before and after leadership studies



**Fig. 6** NOTECHS ratings of control group (dark grey) and experimental group (light grey) participants commanding a boat

received a mean rating of 3.9 in both groups, with  $sd = 0.6$  in the experimental group and  $sd = 0.7$  in the control group,  $t(101) = -0.2$ ,  $p = 0.867$ . Ratings of cooperation skills averaged to 4.2 ( $sd = 0.5$ ) in the control group and 4.1 ( $sd = 0.4$ ) in the experimental group. The difference between the groups was not significant,  $t(101) = 0.5$ ,  $p = 0.651$ .

In the category situation awareness, average ratings were 3.8 ( $sd = 0.9$ ) in the control group and 3.9 ( $sd = 0.6$ ) in the experimental group, which was not significantly different,  $t(100) = -0.6$ ,  $p = 0.583$ . The difference between decision making skills of both groups was on chance level as well,  $t(98) = 0.6$  and  $p = 0.565$ , with mean ratings of 3.8 ( $sd = 0.9$ ) and 3.7 ( $sd = 0.8$ ) in the control group and the experimental group, respectively.

### 3.7 Performance

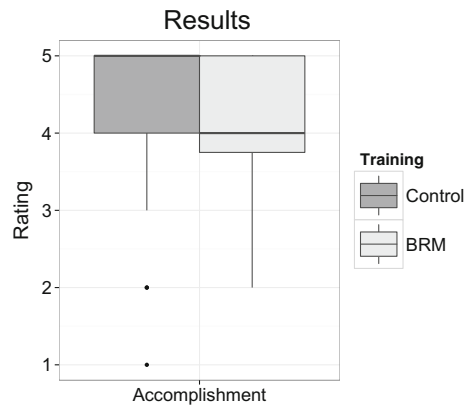
The average rating of task completion achieved by participants in the real-world exercise was 4.3 ( $sd = 1.0$ ) in the control group and 4.1 ( $sd = 0.9$ ) in the experimental group (see Fig. 7). The  $t$  test of this difference did not yield a significant result,  $t(96) = 1.1$ ,  $p = 0.281$ .

## 4 Discussion

A higher effectivity of BRM training as compared to the standard training was found on the level of reactions and learning criteria: the BRM course significantly increased BRM-related knowledge in the experimental group and was rated to be more interesting and relevant as well as better organised and presented than the standard course.

In all of the remaining evaluation criteria, no significant differences between the control group and the experimental group were found. In essence, participants' attitudes regarding communication and coordination, recognition of stressor effects and command responsibility were unaffected by the BRM course. During the nautical

**Fig. 7** Task completion in the real-world exercise by study participants of the control group (dark grey) and the experimental group (light grey)



exercises, participants of the BRM course did not show more effective behaviour as regards leadership, cooperation, decision making or building and sustaining situation awareness. Finally, the BRM course did not lead to better results in terms of task accomplishment during the exercise.

Compared to the results of team training evaluations that were previously reported in the literature (see Salas et al. (2008) for a comprehensive meta-analysis), the effectivity of BRM classroom training was rather low. An influence of team training on attitudes, behaviour and performance of training participants has been repeatedly found in the previous studies, and cognitive effects as reported in Salas et al. (2008) are greater than the knowledge difference observed in the present sample.

This raises the question of why a transfer of the BRM course contents to participants' attitudes, to their behaviour and to the results of the practical exercise could not be found in the present study. The answer may lie in methodological shortcomings of the study or in deficits in the training design.

Regarding the methodology of the study, the application of five different, standardized measures to assess training effects, on each of Kirkpatrick's levels of training evaluation, and in a sample of 117 study participants, at first renders it unlikely that an existing training effect has been missed in data acquisition and analysis. However, there were two aspects of study design that were not under control of the experimenters: First, assignment of the participants to experimental group or control group was not completely randomized, but determined on course level by the German Naval Academy on administrative grounds. This could have led to inequalities between groups already before the training intervention, which could have masked training effects. Analysis of sample equivalence at  $t_0$ , however, showed that no inequalities between groups occurred despite the imperfect way of randomization. And second, the nautical exercises were not exclusively run for data acquisition, but were designed and conducted by instructors of the Leadership Studies for the purpose of training in a realistic setting. This has led to a ceiling effect: even without BRM training, the majority of the participants could fulfil most or even all task requirements, so there was little room left for further improvements in the experimental group (see

Fig. 7). Obviously, the tasks were designed so as to make the goals of the exercises attainable for most junior officers at this stage of their career, which is sensible for training purposes, but at the same time a methodological drawback of the present study.

The absence of training effects on the remaining evaluation criteria cannot be attributed to methodological reasons. Instead, we will attempt to explain the observed results on a general level in the light of research on transfer and, more specifically, based on comparisons with previous studies that did or did not find team training effects.

The aim of the BRM classroom training was to convey general principles of human behaviour, performance, teamwork and leadership that could be applied by participants to various contexts and in various occasions of teamwork. This approach resembles attempts made between the 1950s and the 1980s to train generalizable cognitive skills of analytical thinking by teaching chess, programming or mathematical problem solving. Evaluations of such educational programmes have consistently shown that the skills acquired by students were not applied outside the context in which the skills were taught. In their review on the context-specificity of cognitive skills, Perkins and Salomon (1989) conclude that a transfer of skills and general principles to other contexts rarely occurs spontaneously and is “more a matter of wishful thinking than hard empirical evidence”. This does not mean that such a transfer cannot take place, but it requires deliberate efforts to anticipate and prime the new context and to exercise the application of the skills therein. We believe that this strong context-specificity does not only apply to cognitive skills such as decision making or maintaining situation awareness, but to social skills as for example coordination and communication as well. Therefore, the limited transfer in the present study can be explained with the strong emphasis on generalizable principles in the BRM training at the expense of a specific application of these principles to the context of nautical exercises.

Context-specificity as a determinant of training success is apparent in the comparison of CRM evaluation studies from the field of healthcare as well. Nielsen et al. (2007), for example, conducted a comprehensive randomised controlled trial of the effectiveness of teamwork training for medical personnel in obstetrics. Similar to the BRM training course in our study, the teamwork training course was composed of lectures on CRM-principles and group exercises in the classroom, but contained no exercises in applying these principles during (simulated) labour and delivery care. This training had no impact on clinical outcomes and improved only one out of 11 process indicators of delivery care. In contrast, Thomas et al. (2007) found significant improvements of team processes during simulated neonatal resuscitation for five of six indicators. The focus of this training course was teamwork behaviours in the context of neonatal resuscitation, which were presented in a lecture, discussed with the participants, demonstrated in videos and practiced through role play as well as in simulated resuscitations.

The broad scope of the BRM classroom training and the lack of priming the specific context of application of BRM principles do not only explain the missing effect on participants' behaviour, but the limited cognitive effects as well. Research

in instructional psychology (Reder 1980) as well as models from cognitive psychology (Anderson 1983) show that retention and subsequent retrieval of information depend on the establishment of connections between new information and existing knowledge structures. If abstract information cannot be fit into a conceptual structure immediately, retention in long-term memory is improbable (Reder (1980) pp. 10–11). When receiving instructions on general BRM principles without a specific context of application, training participants may have difficulties to identify conceptual structures in memory to which the new information can be sensibly connected. This problem may have been particularly accentuated in our sample because at the time of our study, the junior officers had only little experience in commanding and navigating a vessel.

The lack of effects on participants' attitudes cannot be explained on the basis of the available data. On the one hand, the positive evaluation of the BRM course's relevance, organisation and presentation speaks in favour of attitude changes in the desired direction, because a positive evaluation of the source of information can increase the probability of attitude changes according to this information (Petty and Cacioppo 1986; Petty and Briñol 2010). On the other hand, there are many more variables potentially influencing the formation and change of attitudes (Ibid.), and we did not collect data that would be indicative of such variables. Nevertheless, a training design focusing on the application of BRM skills may also be more effective in bringing about attitude changes. If a person is required to act in a way that is not consistent with his or her attitudes, e.g. when practicing procedures during training, this can trigger attitude changes that tend to restore the consistency between attitudes and behaviour (so called dissonance reduction (Festinger 1962; Petty and Briñol 2010)).

## 5 Conclusion

The study presented in this article is the first one assessing BRM training effects on reactions, knowledge, attitudes, behaviour and performance in an experimental control group design. The effectiveness of the classroom-based BRM training under study was overall rather low: although the training was positively perceived and considered useful by the participants, it didn't bring about more than a small gain in resource management knowledge. Attitudes and non-technical skills were not affected by the training and no performance improvement could be observed. The tasks in the nautical exercise allowed for an overall high level of performance of all participants, which made it difficult to detect further improvements in the experimental group. But due to the missing effects on attitudes and behaviour, we consider it unlikely that performance improvements would have been observed in more difficult tasks. The reason for the limited effectiveness of the BRM training under study lies in the strong focus on generalizable principles that may be applicable to a broad set of tasks and contexts, at the cost of specific skills, techniques and behaviours and their application in the upcoming nautical exercise.

As has been described in the outset of this article, averages at sea often involve resource management deficits and CRM training in general has been shown to be

effective in many more fields than in aviation. This is why we are convinced that BRM training should and can be effectively applied to seafaring teams as well. Based on the results of the present study and on the plethora of literature from more than two decades of CRM research and practice (e.g. Cannon-Bowers et al. 1995; Salas et al. 1999; Salas et al. 2006a; Flin et al. 2008; Helmreich et al. 1999; Salas et al. 2006b), we would like to make the following recommendations for BRM training design and implementation:

1. In order to be effective, training should be directed at specific behaviours and best practices in a given context of application. This requires the definition of best practices and behavioural standards during training design. General information are an important starting point to justify best practices and to motivate participants to act accordingly, but will not sufficiently transfer into tangible improvements in non-technical skills.
2. Train complete teams instead of individual team members. New practices and procedures are more stable if they have been introduced to and exercised by each member of the working team aboard.
3. Training contents should be tailored to the individual teams. Determine training needs at the beginning of the training and focus the training on those non-technical skills and procedures that do not sufficiently comply with the behavioural standards.
4. Provide opportunities for repeated training and debriefing of behavioural standards. The more often behaviours have been practised, the more likely is their retention and application.
5. Follow a step-by-step approach in training instead of trying to improve everything at once. Focus on no more than three behavioural standards at a time. If a standard is exhibited repeatedly and sufficiently, go on to the next one.
6. In simulators, technical and non-technical skills should be trained jointly, because they must be jointly executed on the bridge. This means that non-technical skills training should not be limited to specific BRM courses. Effective leadership, communication, coordination, decision making and situation awareness should be propagated and encouraged at all occasions of simulator training where they influence the performance in the exercise.

There are two key issues that should be addressed by future research on BRM training. First, it would be worthwhile to conduct research on the effectiveness of BRM training programs that follow the above mentioned principles, thus establishing an empirically based standard for the design and implementation of BRM trainings. Such trainings are required by the STCW convention, alternatively to approved in-service experience, as a means to demonstrate sufficient resource management proficiency. In relation to this, a second research question would be to determine whether BRM trainings and in-service experience are indeed equally effective in achieving sufficient resource management skills. Results of these research endeavours will contribute to further the development of BRM trainings and to disseminate sufficient BRM skills in seafaring teams.

## References

- Alliger GM, Tannenbaum SI, Bennett J, Winston H, Traver H, Shotland A (1997) A meta-analysis of the relations among training criteria. *Pers Psycho* 50(2):341–358
- Anderson JR (1983) A spreading activation theory of memory. *J Verb Learn Verb Be* 22:1–33
- Barnett M, Gatfield D, Pekcan C (2003) A research agenda in maritime crew resource management. In: *Proceedings of the International Conference on Team Resource Management in the 21st Century*
- Brun W, Eid J, Johnsen BH, Laberg JC, Ekornas S, Kobbeltvedt T (2005) Bridge resource management training: Enhancing shared mental models and task performance. In: Montgomery H, Lipshitz R, Brehmer B (eds) *How professionals make decisions*. Lawrence Erlbaum Associates, pp S. 183–193
- Byrdorf P (1998) Human factors and crew resource management: an example of successfully applying the experience from CRM programmes in the aviation world to the maritime world. In: *Proceedings of the 23rd Conference of the European Association for Aviation Psychology*
- Cannon-Bowers JA, Tannenbaum SI, Salas E, Volpe CE (1995) Defining competencies and establishing team training requirements. In: Guzzo R, Salas E (eds) *Team effectiveness and decision-making in organisations*. Jossey-Bass, San Francisco, pp S. 333–380
- Carter-Trahan AC (2009) An examination of the human factors attitudes and knowledge of surface warfare officers. Unveröffentlichte Diplomarbeit, Naval Postgraduate School, Monterey, CA
- 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. *Accid Anal Prev* 59:26–37. doi:10.1016/j.aap.2013.05.006. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0001457513001978>
- European Commission (2012) Commission Regulation No 965 / 2012. Requirements and administrative procedures related to air operations pursuant to Regulation (EC) No 216/2008 of the European Parliament and of the Council. *Off J Eur Union* L296:1–146
- Federal Aviation Administration (ed) (2013) Code of Federal Regulations. Title 14: Aeronautics and Space. Part 61 - Certification: Pilots, flight instructors and ground instructors. Office of the Federal Register of the United States of America. Retrieved from <http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=40760189a03dfea0b501608f33820a45&rgn=div5&view=text&node=14.2.0.1.1.2&idno=14>
- Festinger L (1962) *A theory of cognitive dissonance*. Stanford University Press, Stanford
- Flin R, Martin L, Goeters KM, Hoermann HJ, Amalberti R, Valot C, Nijhuis H (2003) Development of the NOTECHS (non-technical skills) system for assessing pilots CRM skills. *Hum Factors Aerosp Saf* 3(2):97–119
- Flin R, O'Connor P, Crichton M (2008) *Safety at the sharp end. A guide to non-technical skills*. Ashgate, Aldershot, UK
- Fonne V, Fredriksen O (1995) Resource management and crew training for HSV-navigators. In: Jensen R (ed) *Proceedings of the 8th International Symposium of Aviation Psychology*, Columbus, Ohio, pp 585–590
- Gregorich S, Helmreich RL, Wilhelm J (1990) The structure of cockpit management attitudes. *J Appl Psychol* 75(6):682–690
- Helmreich RL (1984) Cockpit management attitudes. *Hum Factors* 26(5):583–589
- Helmreich RL, Merritt AC, Wilhelm JA (1999) The evolution of crew resource management training in commercial aviation. *Int J Aviat Psychol* 9(1):19–32
- Hetherington C, Flin R, Mearns K (2006) Safety in shipping: The human element. *J Safety Res* 37(4):401–411
- Holgado Tello FP, Chacón Moscoso S, Barbero García I, Sanduvete Chaves S (2006) Training satisfaction rating scale. Development of a measurement model using polychoric correlations. *Eur J Psychol Assess* 22(4):268–279
- Hörmann H-J, Maschke P (1991) Exogeneous and endogeneous determinants of cockpit management attitudes. In: *Proceedings of the 6th International Symposium of Aviation Psychology*, pp 384–390
- IMO (2010) *Seafarers' training, certification and watchkeeping code*
- Kirkpatrick DL (1979) Techniques for evaluating training programs. *Train Dev J*, 178–192
- Nielsen PE, Goldman M, Mann S, Shapiro DE, Marcus RG, Pratt SD et al. (2007) Effects of teamwork training on adverse outcomes and process of care in labor and delivery: a randomized controlled trial. *Obstet Gynecol* 109(1):48–55

- O'Connor P (2011) Assessing the effectiveness of bridge resource management training. *Int J Aviat Psychol* 21(4):357–374
- O'Connor P, Campbell J, Newon J, Melton J, Salas E, Wilson K (2008) Crew resource management training effectiveness: a meta-analysis and some critical needs. *Int J Aviat Psychol* 18(4): 353–368
- O'Connor P, Hörmann HJ, Flin R, Lodge M, Goeters K (2002) Developing a method for evaluating crew resource management skills: A European perspective. *Int J Aviat Psychol* 12(3):263–285
- Perkins DN, Salomon G (1989) Are cognitive skills context-bound? *Educ Res* 18(1):16–25
- Petty RE, Briñol P (2010) Attitude change. In: Baumeister RF, Finkel EJ (eds) *Advanced social psychology: The state of the science*. Oxford University Press, London, pp 217–259
- Petty RE, Cacioppo JT (1986) The elaboration likelihood model of persuasion. *Adv Exp Soc Psychol* 19:124–164
- R Development Core Team (2010) R: A language and environment for statistical computing [Software-Handbuch]. Vienna, Austria, Harleben. Retrieved from <http://www.R-project.org> (ISBN 3-900051-07-0)
- Reder LM (1980) The role of elaboration in the comprehension and retention of prose: a critical review. *Rev Educ Res* 50(1):5–53
- Röttger S, Vetter S, Kowalski J (2013) Ship management attitudes and their relation to behaviour and performance. *Hum Factors* 55(3):659–671
- Salas E, DiazGranados D, Klein C, Burke CS, Stagl KC, Goodwin GF et al. (2008) Does team training improve team performance? A meta-analysis. *Hum Factors* 50(6):903–933
- Salas E, Prince C, Bowers CA, Stout RJ, Oser RL, Cannon-Bowers JA (1999) A methodology for enhancing crew resource management training. *Hum Factors* 41(1):161–172
- Salas E, Wilson KA, Burke CS, Wightman DC, Howse WR (2006a) A checklist for crew resource management training. *Ergon Des* 14(6):6–15
- Salas E, Wilson KA, Burke CS, Wightman DC, Howse WR (2006b) Crew resource management training research, practice, and lessons learned. *Rev Hum Factors Ergon* 2(1):35–73
- Staufenbiel T (2000) Fragebogen zur evaluation von universitären Lehrveranstaltungen durch studierende und lehrende. *Diagnostica* 46(4):169–181
- Thomas EJ, Taggart B, Crandell S, Lasky RE, Williams AL, Love LJ et al. (2007) Teaching teamwork during the neonatal resuscitation program: a randomized trial. *J Perinatol* 27(7):409–414
- van Avermaete JAG, Kruijsen EAC (1998) NOTECHS: The evaluation of non-technical skills of multi-pilot aircrew in relation to the JAR-FCL requirements (Bericht). NOTECHS Consortium