The Journal of Navigation (2021), **74**:5 1127–1141
doi:10.1017/S0373463321000436

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Experimental study on the effects of a single simulator-based bridge resource management unit on attitudes, behaviour and performance

Stefan Röttger,* and Hannes Krey

Department of Naval Psychology and Ergonomics, Naval Institute of Maritime Medicine, Kronshagen, Germany.
*Corresponding author. E-mail: stefan.roettger@gmx.de

Received: 22 February 2021; **Accepted:** 25 April 2021; **First published online:** 17 May 2021

Keywords: bridge resource management, simulator training, training effectiveness, experimental study

Abstract
The objective of this work was to assess whether the implementation of a bridge resource management (BRM) unit into the simulator-based nautical training of the German Navy is effective in improving non-technical skills and navigation performance. To this end, questionnaire data, observations of behaviour and performance outcomes were compared between a control group and an experimental group. Data of 24 bridge teams (126 sailors) were used for the analyses. Ten teams received BRM training and 14 teams served as the control group with unchanged simulator training. Reactions to simulator training were positive in both groups but more favourable in the control group. In the BRM group, significantly more positive attitudes towards open communication and coordination, more frequent sharing of information and fewer collisions were found than in the control group. Effect sizes were rather small. This may be due to the limited scale of the BRM unit, which consisted of only one instruction-training-feedback cycle. The extension of BRM-related feedback to all simulator runs of the nautical training can be expected to produce larger effects on attitudes, behaviour and performance.

1. Introduction

Safety in shipping is affected by a multitude of factors, among them the non-technical skills of the crew (Hetherington et al., 2006). These skills encompass cognitive and interpersonal competencies that are necessary for safe navigation and ship operation by the crew. Examples are team communication, decision making, situation awareness, leadership or managerial skills to effectively use all available technical and personal resources during routine operations and in emergencies (Wahl and Kongsvik, 2018). Training to improve such non-technical skills, called crew resource management training (CRM), was first developed in commercial aviation in response to a series of fatal flight accidents in the 1970s (Helmreich et al., 1999). In seafaring, resource management, leadership, teamwork and managerial skills were incorporated into the Standards of Training, Certification and Watchkeeping (STCW) of the International Maritime Organization (IMO) with the 2010 Manila Amendments. This has led to a marked increase in research and development of maritime resource management training. Such training is required for the nautical department, called bridge resource management training (BRM), as well as for the engineering department, for which the term engine room resource management training (ERM) was established. One theme in the maritime resource management literature is the demand for practice (e.g. Barić et al., 2018; Sanjeev Vakil, 2019; Cavaleiro et al., 2020). Although the STCW do not prescribe a certain training method, the common line of argument is that trainees must experience and perform teamwork and leadership behaviours in order to develop the required skills. Accordingly, studies on the effectiveness of classroom-based resource management training so far show hardly any transfer to applications in the simulator (Saeed et al., 2017) or in the real world (Röttger et al., 2016). While the use of simulators for resource management training seems to be more promising, designing and implementing such training may be difficult for two reasons. Firstly, simulator time is a costly and limited resource, which is not easily available. Secondly, there is still too little evidence on effective simulator-based resource management training that could guide training design and implementation and justify the more expensive use of simulators. The review of Havinga et al. (2017) gives a good impression of the state of affairs in research on maritime resource management training up to 2015. The authors found eight studies reporting formal outcome measures of maritime resource management training, of which six employed simulators. Because of methodological shortcomings, not one of these studies allowed for a valid conclusion on the effectiveness of the training. Most severe methodological problems were lack of a control group or of a control condition, missing statistical analysis to determine whether observed effects exceed measurement error, or small sample sizes insufficient to detect systematic training effects.

From 2016 on, two further publications showed that simulation-based maritime resource management training is positively perceived and evaluated by training participants (Hong and Kim, 2016; Espevik et al., 2017). Again, actual training effects in terms of improvements of knowledge, behaviour or performance were not in the scope of these studies.

The first and, up to now, only experimental study of such training effects was published by Tvedt and colleagues in 2018. In this study, BRM-related knowledge and attitudes as well as behaviour and performance in simulator scenarios were assessed in 94 bridge officers before and after their participation in BRM training. The training started with two days of classroom lectures, followed by two days of simulator exercises. Only one simulator exercise served for actual training, i.e., giving feedback on teamwork and situation awareness. The remaining simulator time had to be used for familiarisation, pre-test and post-test measurements. Tvedt et al. (2018) found that the BRM course was positively evaluated by training participants, and that BRM-related knowledge as well as some BRM-related attitudes had significantly improved. Observer ratings of teamwork, situation awareness and mission success in the simulator, however, did not improve over the course of the BRM training. The missing training effects on behaviour and performance may reflect the fact that with only one feedback session, a rather small part of the overall training time was devoted to directly addressing the actions and procedures of the bridge team. Another explanation could be that the training participants formed synthetic teams, i.e., they did not work together as a team on a regular basis. Thus establishing a common way of working together may have interfered with adopting the teamwork behaviours taught in the BRM training.

2. Research question

The purpose of this study was to determine whether including a BRM training unit into the nautical team training of the German Navy improves BRM-related attitudes, behaviour and performance of the training participants. The standard nautical team training is a nautical simulator training for bridge teams with a focus on ship handling and navigation. Bridge teams book such training, for example, to prepare for an upcoming voyage or to train for manoeuvres in advance of a real-life exercise. The purpose of including a BRM unit was (1) to additionally address and increase resource management skills during this training, while (2) providing a BRM training opportunity to operational bridge teams that have a tight schedule and could not accommodate the regular three- or four-day BRM training.

In doing so, this study goes beyond the study of Tvedt et al. (2018) in three ways. First, natural bridge teams are studied, i.e., teams of seafarers who work together on the bridge of their ship on a daily basis. Thus, it is assumed that the basic processes of working together were clear, so participants were better able to focus on implementing the resource management skills introduced and debriefed by the instructors. Second, in order to capture training effects on behaviour reliably, video recordings of simulator scenarios were analysed with a detailed and exhaustive coding manual and all observable behaviours were counted. Third, a control group design was employed, thus it was possible to ascertain

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

the effect of BRM-enriched simulator training compared with conventional nautical simulator training of bridge teams.

3. Methods

3.1. Participants

Thirty bridge teams from 30 different vessels of the German Navy participated in the study. Due to malfunctions of audio and video recordings, data from six teams could not be used for analysis. This affected five data sets in the BRM group and one data set in the control group. The remaining sample consists of 24 bridge teams with 126 sailors of the German Navy. Training participants formed natural teams, i.e., they worked together on the bridge of their respective ships. Fourteen teams (72 sailors, 22% females) served as control group. Relative frequencies of military ranks in this group are: 54% officers, 28% petty officers and 18% enlisted ranks. Ten teams (54 sailors) were in the experimental group and received the BRM unit. This group consisted of 11% females, 54% officers, 35% petty officers and 11% enlisted ranks. No other personal or demographic information was collected to ensure the anonymity of participants in the data analyses. BRM group and control group were opportunity samples, i.e., the bridge teams visited the simulator facility for nautical training and volunteered to participate in the data collection and the BRM training, respectively. Bridge teams participating in the study had not received simulator-based BRM training before. Each team consisted of at least five sailors: commander, officer of the watch, helmsman, radar operator and ECDIS (Electronic Chart Display and Information System) operator. Occasionally, a sixth sailor accompanied the teams. His or her task was to support either the radar or the ECDIS operation, or to operate the radio set.

3.2. BRM training unit

The structure of the training unit as administered to the BRM group is depicted in Figure 1. Training commenced with an introductory lecture on the reasons for conducting BRM training and on principles of effective non-technical skills. More specifically, leadership and management, cooperation in a team, building and maintaining situation awareness, decision making and effective communication were addressed. These topics were illustrated by examples of successful and unsuccessful teamwork or leadership behaviours in navigation.

The lecture was followed by a one-hour simulator exercise (Figure 2). The task of the bridge teams was to apply the non-technical skills while navigating into Eckernförde Bay and mooring in the naval harbour. The demand for effective teamwork and leadership was increased by dense traffic, rough weather with impaired visibility, and areas with restricted room to manoeuvre. After the exercise, the bridge team received a video-aided debriefing regarding the successful application of non-technical skills and the possibilities for further improvements. In order to achieve detailed and complete feedback, three instructors were involved in observing and debriefing the five to six sailors of each bridge team.

3.3. Measures

Differences between control group and BRM group were assessed on four levels of training evaluation as proposed by Kirkpatrick (1979). These are (1) the evaluation of the training by the participants, called reactions, (2) learning effects in terms of cognitive changes as, for example, in knowledge or attitudes towards the training contents, (3) behaviour of training participants while performing the trained tasks, and (4) results of performing the trained behaviours, for example, in terms of quality, quantity, efficiency or safety. Although Kirkpatrick's original assumption, that lower level training effects are the best predictor and a prerequisite for effects on the next higher level, did not hold empirically (Alliger et al., 1997), the four levels are a useful systematic to guide the design of a comprehensive training evaluation. The individual measures are described in the following paragraphs.

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

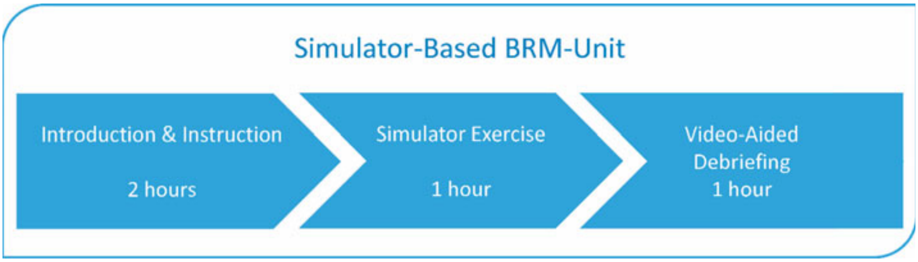

Figure 1. *Structure of the BRM training unit.*

Figure 2. *Bridge team during simulator exercise. One of the cameras used for video-aided debriefing and data collection can be seen in the forward corner on starboard.*

3.3.1. Reactions

Participants' reactions were assessed with a nine-item questionnaire (see Table 1) based on the results of Staufenbiel (2000) and Holgado Telo et al. (2006). They found that course evaluations as given by learners usually refer to three different aspects: (a) organisation and presentation of course contents, (b) interest of the contents and their relevance for the learners' work activities, and (c) a more global, affective evaluation. The scale ranged from one to five with values below three indicating rather negative evaluations and values above three rather positive evaluations.

3.3.2. Learning

Cognitive effects were assessed in terms of attitude changes. In the early days of CRM in aviation, studies had shown that many pilots held dysfunctional attitudes about, for example, how a cockpit team should communicate, who has to take responsibility for the flight's safety, or how stress and fatigue affect pilots' capabilities (e.g. Foushee, 1984; Helmreich, 1984). Using the standardised Cockpit Management Attitudes Questionnaire (CMAQ), Helmreich et al. (1986) could correctly classify the performance ratings of 96% of the pilots in their study as being below or above average. Thus, attitude changes are a

Table 1. *Scales and items of the 'Reactions' questionnaire.*

Organisation and presentation	Explanations and instructions were clear and easy to understand.
	Comments and hints from trainers regarding performance of participants were helpful.
	Comments and hints from trainers regarding behaviour of participants were helpful.
Interest and relevance	Course contents are useful for my duties.
	I could improve my nautical skills in this course.
	I could improve my teamwork skills in this course.
	I can apply what I have learned in this course.
Global	I liked the course.
Evaluation	(–) I would assign the course the following overall grade (1: very good to 5: very bad)

Note: (–) indicates negative polarity.

viable training aim in their own regard, and resource management attitudes were assessed with the Ship Management Attitudes Questionnaire – German Navy (SMAQ-GN, Röttger et al., 2013), an adaptation of the CMAQ for use in naval seafaring. The questionnaire consists of three scales that cover attitudes towards communication and coordination, command responsibility and recognition of stress effects on a scale ranging from one to five. Higher values indicate more effective attitudes for BRM.

3.3.3. Behaviour

Behaviour of study participants was recorded with three video cameras (see Figure 3) for subsequent offline-review with Biobserve Spectator, a software solution for observational data acquisition and analysis.

To minimise bias in the video analyses, two measures were taken. First, the analyst was blind towards the experimental condition (Hróbjartsson et al., 2013), i.e., they did not know whether the bridge teams belonged to the control group or to the BRM group. Second, analysis was based on registering observable behaviours instead of judging the participants' performance (Kent and Foster, 1977, 283–284; Hartmann and Wood, 1990, 118–119). To this end, the Exhaustive Bridge Team Interaction Coding Scheme (EBTICS, Röttger et al., 2015) was developed, a coding manual designed to be exhaustive for any act of communication of a navigating bridge team. Table 2 gives an overview of the main categories of EBTICS. The complete coding scheme can be obtained from the corresponding author or downloaded from researchgate.net.

The main categories of EBTICS reflect common CRM/BRM principles, as described, for example, in the NOTECHS taxonomy (van Avermaete and Kruijsen, 1998; Flin et al., 2003). For building and maintaining a shared situation awareness, the relevant behaviours are: communicating information on the own ship and the environment as well as discussing situation assessments. Techniques of effective leadership as defined in NOTECHS are: statement of intentions and plans, creating an open team atmosphere by giving positive personal feedback and asking for or considering suggestions of the crew. The category suggestions also pertains to decision making and speaking-up behaviour, the latter being specifically targeted by some BRM courses (Espevik et al., 2017; Barić et al., 2018). Cooperation and mutual support, at the very core of teamwork, are mirrored in the EBTICS categories of giving personal feedback, providing assistance or support, and communicating about personal conditions of teammates. Finally, techniques to ensure proper communication, which are not contained in NOTECHS, were included. For example repeating/confirming orders and reports or making sure that someone is actually attending to a message.

In contrast to the assessment systems and rating scales for BRM training that have been published so far (e.g., Saeed et al., 2017; Bolstad, 2018; Tvedt et al., 2018; Da Conceição et al., 2019), EBTICS

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

Table 2. *Overview of the EBTICS categories for analysis of participants' teamwork behaviour.*

Main category	Description
Order/Instruction	Give/repeat/check back on order, confirm repeated order, report execution.
Objective/Intention/Plan	Ask for/state/repeat/check back on intentions and plans. (e.g., 'I want to pass the danger buoy on the port side').
Information ship	Ask for/give/confirm information on own ship (e.g., position, radar settings, status of moorings).
Information environment	Ask for/give/confirm information on the environment (e.g., weather, traffic, depth).
Situation assessment	Ask for/give/confirm information on current situation assessment in terms of conclusions from available information on ship and environment (e.g., sufficient room for a manoeuvre, risk of collision).
Suggestion	Ask for/give/evaluate a proposed course of action.
Personal feedback	Utterances relating to interpersonal relations/atmosphere in the team (e.g., to thank, criticise, encourage a team member).
Personal condition	Ask for or state personal condition (e.g. being unwell, tired or nervous).
Assistance/Support	Offer/provide/reject/ask for help in performing a task or solving a problem.
Establishing contact	Getting someone's attention/making sure an addressed person takes notice/making clear whether one can attend to someone else at the moment.

was developed for research purposes with an emphasis on behaviour definitions that refer to directly observable targets, are easy to understand and require as little interpretation as possible (Hartmann and Wood, 1990, 109–110). This fosters objectivity, but makes the coding of observations too fine-grained and time-consuming for application in a training setting. Data from a second analyst who independently coded video recordings of 16 participants showed an inter-rater-reliability of .74 (Röttger et al., 2015).

Each act of verbal and non-verbal communication of the commander, officer of the watch, helmsman, radar operator and ECDIS operator was classified with EBTICS and counted. As the absolute number of observed behaviours varied with the duration of the simulator run, relative frequencies (utterances per minute) entered the statistical analysis.

3.3.4. Results

To evaluate the impact of the training in terms of performance outcome, each bridge team was confronted in the test scenario with a vessel that was on collision course but difficult to detect. The result of this critical situation could be (1) timely detection of the vessel without necessity of a last-minute manoeuvre; (2) avoiding collision with last-minute manoeuvre (which can be considered a near miss); (3) collision despite last-minute manoeuvre; (4) collision without attempting last-minute manoeuvre. Performance outcome was assessed for each bridge team, and frequencies of collisions and last-minute manoeuvres were compared between BRM group and control group.

3.4. Procedure

Figure 4 gives an overview of the data acquisition procedure for the BRM group and control group. The study was conducted in the ship's bridge simulator of the German Navy as part of the standard nautical simulator training of bridge teams. This training lasts between two and three days depending on the training needs and the availability of the bridge team in question. At the start of the first day of simulator

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

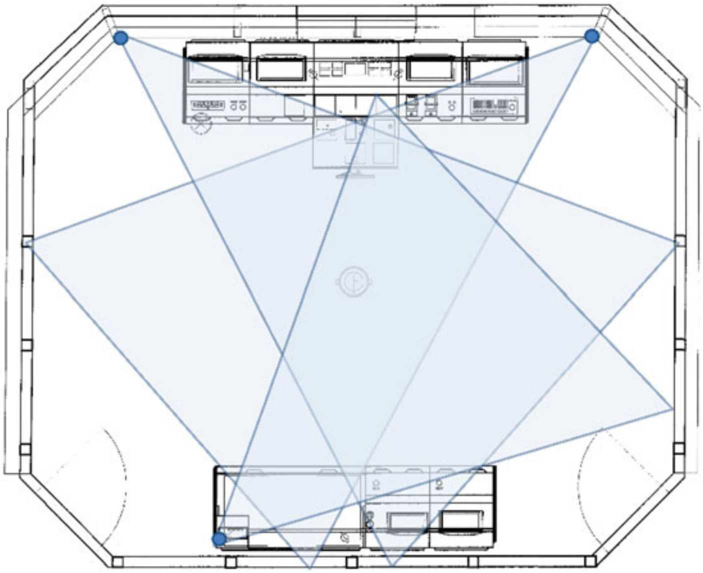

Figure 3. Layout of the ship's bridge simulator with camera positions (dots) and areas covered by the cameras (triangles).

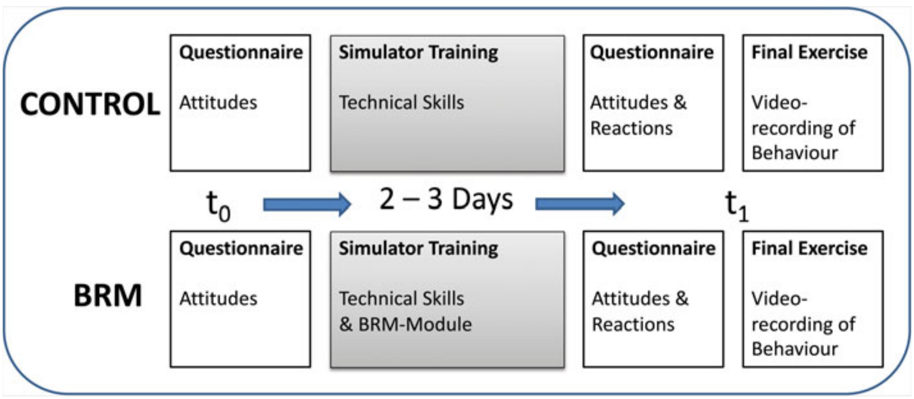

Figure 4. Overview of study procedure in experimental group (BRM) and control group.

training (t1), all participants signed informed consent and the questionnaire on ship management attitudes was administered. Then, the control group went through the traditional simulator training with an emphasis on technical and navigational skills. The BRM group received the BRM module during their simulator training. The time needed for the BRM module was compensated for by shortening the training time for nautical simulator exercises. At the end of the last day of their simulator training (t2), all participants filled in the reactions questionnaire and the attitudes questionnaire and went through the test scenario, which was video recorded for later analysis of behaviour and performance.

3.5. Statistical analysis

Analyses were conducted with the software *R* version 3.4.2 (R Core Team, 2017). Means (M) and standard deviations (SD) of measurements are reported in the text and tables. In figures, standard

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

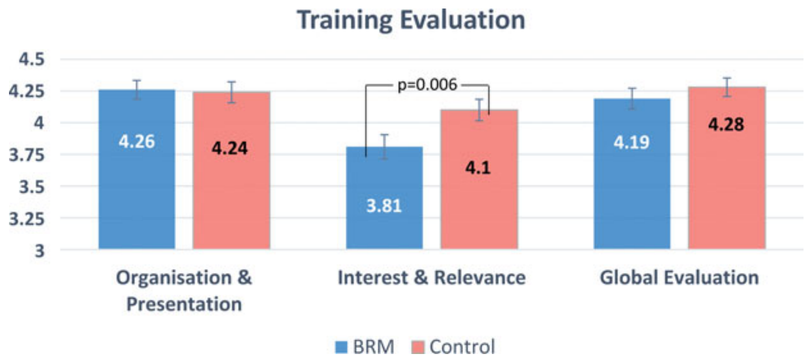

Figure 5. *Means and standard errors of training evaluations in BRM group and control group.*

error is used as a measure of dispersion. Differences in attitudes were tested with *t*-tests for dependent samples (comparison of pre- and post-training values within each group) and for independent samples (comparison between control group and BRM group at the end of the training). Scores of the reaction questionnaire and the behaviour frequencies did not meet the distribution requirements for t-tests, so the Wilcoxon rank-sum test was used instead to analyse differences between the control group and the BRM group. Pearson's χ^2 test was performed to compare the frequencies of all performance outcomes. Statistical significance of a difference in collision frequencies only could be determined with the more sensitive Lancaster's mid-*P* value (Lancaster, 1961; Hwang and Yang, 2001). Statistical tests of differences in participants' reactions was based on the assumption that simulator training including a BRM unit could be evaluated as better than, as well as worse than, traditional simulator training. Thus, two-sided significance tests were conducted. Regarding ship management attitudes, behaviour and performance, it was assumed that the training would have either a positive effect or no effect at all, because negative effects on these outcome variables have not been reported in the CRM literature (e.g. Helmreich and Wilhelm, 1991; Salas et al., 2001; O'Connor et al., 2008; Marquardt et al., 2010). Accordingly, significance tests for these training effects were one-sided.

4. Results

4.1. Training evaluation

Means and standard errors of training evaluations in control group and BRM group are depicted in Figure 5. With values greater than three on a five-point scale, evaluations were consistently positive. The means of global evaluation were 4·19 (SD = 0·61) in the BRM group and 4·28 (SD = 0·59) in the control group and did not differ significantly, W(54,72) = 2117·5, $P = 0·375$. Organisation and presentation of training contents, too, were evaluated equally well in both groups (BRM group 4·26, SD = 0·62; control group 4·24, SD = 0·61), W(54,72) = 1895, $P = 0·808$. With a mean of 3·75 (SD = 0·68), interest and relevance of the simulator training was perceived significantly lower in the BRM group than in the control group, which gave a mean rating of 4·25 (SD = 0·77), W(54,72) = 2493·5, $P = 0·006$.

4.2. Ship management attitudes

Figure 6 gives an overview of means and standard errors of the SMAQ-GN scales in both groups before and after simulator training. Ship management attitudes did not differ significantly between BRM group and control group before the start of the simulator training. The means of the communication and coordination scale were 4·04 (SD = 0·41) in the BRM group and 4·07 (SD = 0·37) in the control group, t(117) = 0·46, $P = 0·644$. The command responsibility scale showed mean values of 2·82 (SD = 0·7) and 2·74 (SD = 0·64) in BRM group and control group, respectively, t(117) = –0·66, $P = 0·51$. The

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

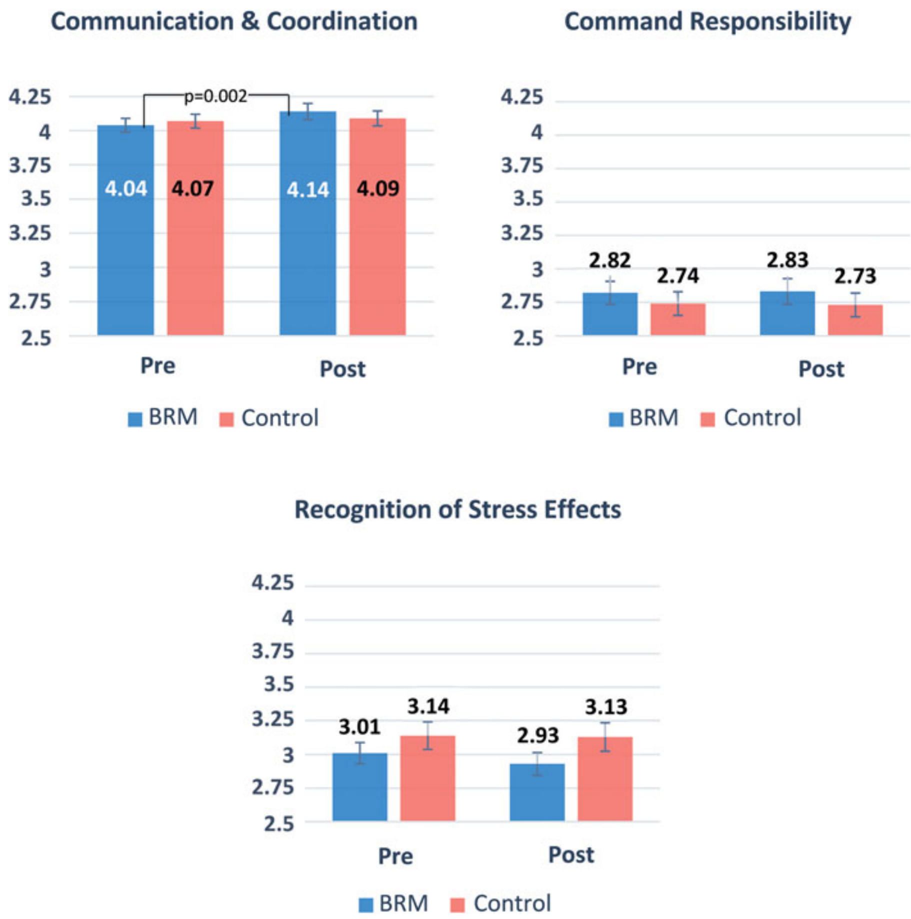

Figure 6. *Means and standard errors of ship management attitudes in BRM group and control group before and after simulator training.*

means of the scale regarding recognition of stress effects were 3·01 (SD = 0·64) and 3·14 (SD = 0·74), t(117) = 0·98, $P = 0{\cdot}328$.

The hypothesis that the BRM group would show more effective ship management attitudes than the control group after the training was not confirmed. Statistical analysis yielded no significant differences in the direction of higher scale values in the BRM group compared with the control group after training. The mean values in for the communication and coordination scale were 4·14 (SD = 0·44) and 4·09 (SD = 0·45) for BRM and control, t(120) = −0·58, $P = 0{\cdot}28$. The mean values for the command responsibility scale were 2·83 (SD = 0·71) and 2·73 (SD = 0·73) in BRM group and control group, t(120) = −0·73, $P = 0{\cdot}232$. The recognition of stress effects scale did not show a higher mean value in the BRM group (2·93, SD = 0·62) compared with the control group (3·13, SD = 0·87) either, t(120) = 1·43, $P = 0{\cdot}922$.

However, statistical analysis of differences between pre- and post-training measurements, and thus changes of participants' ship management attitudes over the course of the simulator training, showed a significant increase in the communication and coordination scale in the BRM group, t(52) = −3·07, $P = 0{\cdot}003$. No such attitude change was detected in the control group, t(61) = 0·04, $P = 0{\cdot}516$. No other increases in effective ship management attitudes were found in the control group [command responsibility: t(61) = 0·68, $P = 0{\cdot}75$; recognition of stress effects: t(61) = 0·61, $P = 0{\cdot}729$] or the BRM group [command responsibility: t(52) = −0·08, $P = 0{\cdot}47$; recognition of stress effects: t(52) = 1·1, $P = 0{\cdot}862$].

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

Table 3. *Mean frequencies (counts per minute) of individual communication and teamwork behaviours in BRM group and control group during final simulator run.*

	BRM		Control			
	M	SD	M	SD	W	*P*
Order/Instruction	0·89	1·1	0·93	1·01	2109	0·793
Objective/Intention/Plan	0·12	0·15	0·15	0·2	2099	0·783
Information ship	0·18	0·17	0·28	0·28	2249	0·934
Information environment	0·59	0·41	0·48	0·36	1639·5	0·067
Situation assessment	0·11	0·13	0·05	0·07	1417	0·004
Suggestion	0·07	0·09	0·06	0·07	1908·5	0·429
Personal feedback	0·01	0·02	0·02	0·05	2045	0·751
Personal condition	<0·01	0·01	0	–	–	–
Assistance/Support	<0·01	0·01	0·01	0·03	2334·5	0·997
Establishing contact	0·25	0·25	0·23	0·23	1846·5	0·316

Note: M: mean. SD: standard deviation. W: Wilcoxon test statistic. p: significance test of hypothesis BRM > control. For descriptions of the behaviour categories, see Table 2.

4.3. Behaviour

Table 3 shows the mean frequencies of individual communication and teamwork behaviours in the BRM group and the control group during the final simulator run, along with statistical test results. Giving and confirming orders and reporting their execution was the most frequent teamwork behaviour, and was not carried out more often in the BRM group than in the control group. Communicating about the external environment was the second-most frequent teamwork behaviour. This was observed more often in the BRM group. Wilcoxon test showed that this difference closely missed statistical significance. Establishing/assuring contact with a teammate and communicating about the own ship ranked third in the mean frequencies of individual teamwork behaviours and did not show a better performance in the BRM group. Communicating about intentions and plans did not show an advantage of the BRM group either. Communications about the current nautical situation, however, occurred twice as frequently in the BRM group as in the control group, which is statistically a highly significant difference. No such difference was found in the frequencies of suggestions.

The remaining teamwork behaviours of giving personal feedback, informing or asking about personal conditions, and asking for or giving support occurred too infrequently during the final simulator run to base any conclusion on these observations.

Restricting the analysis of information exchange to those members of the bridge team that work together most closely and frequently in navigation, i.e., officer of the watch, radar operator and ECDIS operator, gives the results depicted in Table 4. As can be expected, the mean frequencies are higher. The surplus of the BRM group in communicating about the environment and about situation assessments becomes larger and is statistically significant. Figure 7 shows the distribution of the individual frequencies in exchanging information about the aforementioned issues in both groups. The pattern in the histograms suggests that the observed mean differences are due to the fact that there are fewer training participants in the BRM group that communicate rarely or not at all about the environment and situation assessments.

4.4. Performance

Table 5 shows the number of bridge teams that took proper evasive actions in due time (no collision, no last-minute manoeuvre), that encountered a near miss (no collision, last-minute manoeuvre), that

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

Table 4. *Mean frequencies (counts per minute) of individual information exchange by officer of the watch, radar operator and ECDIS operator in BRM group and control group.*

	BRM		Control			
	M	SD	M	SD	W	*P*
Information ship	0·24	0·19	0·37	0·32	706·5	0·930
Information environment	0·82	0·34	0·60	0·32	367	0·004
Situation assessment	0·17	0·14	0·08	0·08	334	0·001

Note: M: mean. SD: standard deviation. W: Wilcoxon test statistic. p: significance test of hypothesis BRM > control. For descriptions of the behaviour categories, see Table 2.

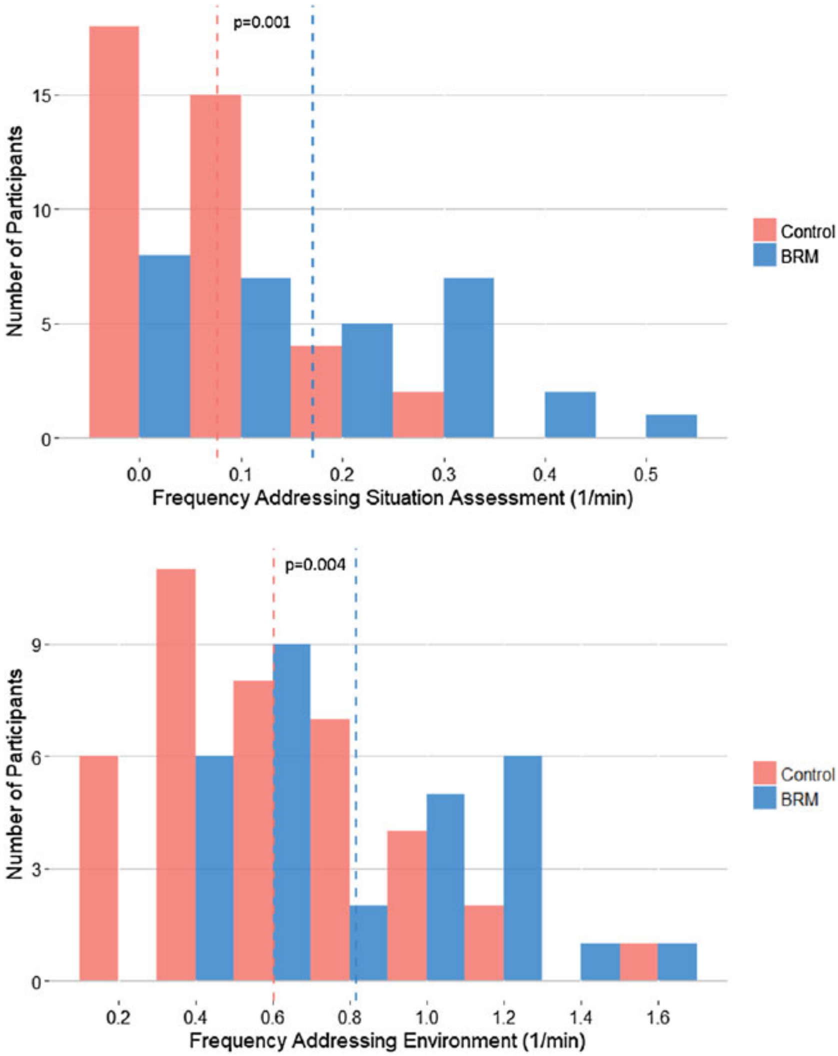

Figure 7. *Individual frequencies of addressing the current situation assessment (upper graph) and the environment (lower graph) by officer of the watch, radar operator and ECDIS operator in BRM group and control group.*

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

Table 5. *Performance in evading a vessel on collision course.*

	No collision no LMM	No collision but LMM	Collision despite LMM	Collision without LMM
BRM	5	5	0	0
Control	5	5	2	2

Note: LMM: Last-minute manoeuvre. $\chi^2 = 3{\cdot}43$, $P = {\cdot}056$.

Table 6. *Number of simulator runs with and without collision in BRM group and control group.*

	Collision	No collision
BRM	0	10
Control	4	10

Note: Lancaster mid-$P = 0{\cdot}047$.

detected the other vessel too late but took evasive actions to reduce the consequences of the upcoming collision (collision despite last-minute manoeuvre) and that became aware of the other vessel too late to attempt any evasive action (collision, no last-minute manoeuvre).

The same number of teams was observed in the BRM and control groups avoiding a collision with or without a last-minute manoeuvre. Four collisions were observed in the control group, half of them without any previous attempt to take evasive action. In the statistical analysis of the joint frequencies of last-minute manoeuvres and collisions, statistical significance is narrowly missed with $\chi^2 = 3{\cdot}43$, $P = {\cdot}056$.

Significance test of the collision frequencies only as depicted in Table 6 gives a Lancaster's mid-*P* value of 0·047. The risk of a collision can be considered significantly lower in the BRM group compared with the control group.

5. Discussion

Significant effects of a single simulator-based BRM training unit on training evaluations, attitudes, behaviour and performance of natural bridge teams were found. This is the first empirical study to show such effects in a control group design and with a behaviour analysis that was blind towards the completion of a BRM training.

In the light of consistent evidence that BRM training is positively perceived by training participants (e.g. Hong and Kim, 2016; Espevik et al., 2017; Havinga et al., 2017; Tvedt et al., 2018), the partly better evaluation of the simulator training without a BRM module may be surprising. This difference can be explained by the fact that bridge teams of the German Navy enter their simulator training with specific training plans. In order to allocate training time to the BRM module, some of these plans had to be skipped. This could have led to lower ratings on interest and relevance of the simulator training in the BRM group compared with the control group. Nevertheless, the ratings from the BRM group were still in the positive range.

An improvement in BRM-related attitudes over the course of the simulator training occurred in the experimental group only. The pattern of attitude change, with a significant increase in positive attitudes towards open communication, but not towards command responsibility and recognition of stress effects, reflects the focus of the BRM module during instruction and feedback. A similar change in attitudes did not occur but could have been expected in the study of Tvedt et al. (2018). However, the attitudes of their training participants towards communication and cooperation were already very positive before the beginning of the training, so a further improvement was hardly possible. The training of Tvedt et al.

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

(2018) did succeed in improving attitudes towards the recognition of stress effects, which demonstrates the potential of combining classroom lectures and simulations to change attitudes of training participants.

Analysis of behaviour counts showed that the bridge teams that had received a BRM training unit did communicate more frequently than the control group about the environment and about assessments of the current nautical situation. This difference is most pronounced and highly significant if the analysis is focused on the officer of the watch, the radar operator and the ECDIS operator. The magnitude and the meaning of the observed frequency differences become apparent if they are translated into the periods of time between two communications. While each of the above mentioned bridge team members in the BRM group communicated on average every 1 min 13 s about the environment and every 5 min 53 s about the current nautical situation, these times were 1 min 40 s and 12 min 30 s in the control group, respectively. The distribution of the communication frequencies suggests that this difference is due to a much lower number of training participants in the BRM group who never, or very rarely, communicate about the environment and the situation assessment.

This advantage of the experimental group over the control group in sharing information and situation assessments may have contributed to the successful separation from a conflicting vessel by all bridge teams of the BRM group in the final simulator scenario. In contrast, bridge teams of the control group performed overall significantly worse, accounting for all four collisions observed in this study.

Although it could be shown that the BRM training module under study successfully improved the attitudes, behaviour and performance of natural bridge teams, it has to be noted that these effects are overall rather small or limited to only a few aspects of attitudes and behaviour. This can be explained by the limited scale of the BRM unit, which lasted just four hours and comprised only one cycle of instruction, exercise and feedback. If the instructions on non-technical skills are scheduled at the beginning of a nautical simulator training, and BRM feedback is provided together with nautical feedback at the end of each simulator run over the course of two to four days, more extensive effects would be expected than those found in this study.

The need for more exercise and feedback opportunities is one conclusion that can be derived from the results presented here and in Tvedt et al. (2018). Another is that it might be easier for natural teams to adopt and apply resource management techniques than it is for teams that work together for the duration of the training only. This observation supports the proposition of Wahl and Kongsvik (2018) to train seafarers within their communities of practice.

Other factors that might be beneficial for resource management training are nautical experience of training participants, blending nautical and resource management training, and adjusting training contents to the individual needs of the training participants. Regarding previous nautical experience, Tvedt et al. (2018) found stronger effects of two days of classroom lectures on knowledge and attitudes of experienced nautical officers than Röttger et al. (2016) found after five days of classroom-based BRM training with junior naval officers. Similarly, resource management training evaluations seem to be more positive among more experienced training participants compared with those with a shorter work experience (e.g. Hong and Kim, 2016). Combining nautical training and resource management training is advocated by Sellberg and Viktorelius (2020) as well as Sanjeev Vakil (2019), for example. They argue that nautical skills and resource management skills should not be treated as separated, because they must be applied in combination and consequently must be trained jointly. Finally, Cavaleiro et al. (2020) stress that BRM training cannot be thought of a one size fits all prescription, and that it has to consider the training needs of the individuals and the teams at hand. Such training needs analysis are commonly called for in team training guidelines, as for example in Salas et al. (2015).

6. Conclusion

With regard to the research question, this study found that including a BRM training unit in the nautical team training of the German Navy does improve BRM-related attitudes, behaviour and performance of the training participants. Regarding the scientific literature, this is the first time that BRM training-related improvements in behaviour and performance have been demonstrated. The methods used in

https://doi.org/10.1017/S0373463321000436 Published online by Cambridge University Press

this study proved to be overall useful to assess training effectiveness. However, future studies should assess more demographic data in order to ensure better comparability of experimental group and control group. In addition, performance data should be assessed in a more comprehensive manner. Open research questions that should be addressed systematically are training effects over a longer period and the hypotheses that BRM training is most effective if conducted with natural teams, integrated into the nautical simulator training, and focused on the training needs of the team at hand.

Author Note. Hannes Krey is now working at LKM GmbH, Kiel, Germany.

References

Alliger, G. M., Tannenbaum, S. I., Bennett, W. Jr, Traver, H. and Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, **50**(2), 341–358.

Barić, M., Čulin, J. and Bielić, T. (2018). Problems that occur in a team: Learning from maritime accidents via simulation training. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, **12**(4), 709–713.

Bolstad, M. (2018). *Virker crew resource management trening? En studie av sjøforsvarets* CRM *kurs' effekt på elever ved skolen*. Master thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Cavaleiro, S. C., Gomes, C. and Lopes, M. P. (2020). Bridge resource management: Training for the minimisation of human error in the military naval context. *The Journal of Navigation*, **73**(5), 1146–1158.

Da Conceição, V. P., Mendes, J. B., Teodoro, M. F. and Dahlman, J. (2019). Validation of a behavioral marker system for rating cadet's non-technical skills. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, **13**(1), 89–96.

Espevik, R., Saus, E. R. and 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., Martin, L., Goeters, K. M., Hoermann, H. J., Amalberti, R., Valot, C. and Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Human Factors and Aerospace Safety*, **3**, 97–120.

Foushee, H. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist*, **39**, 885–893.

Hartmann, D. P. and Wood, D. D. (1990). Observational methods. In: Bellack, A. S., Hersen, M. and Kazdin, A. E. (eds.). *International Handbook of Behavior Modification and Therapy*, Boston, MA: Springer, 107–138.

Havinga, J., De Boer, R. J., Rae, A. and Dekker, S. W. (2017). How did crew resource management take-off outside of the cockpit? A systematic review of how crew resource management training is conceptualised and evaluated for non-pilots. *Safety*, **3**(4), 26.

Helmreich, R. L. (1984). Cockpit management attitudes. *Human Factors*, **26**, 583–589.

Helmreich, R. L. and Wilhelm, J. A. (1991). Outcomes of crew resource management training. *The International Journal of Aviation Psychology*, **1**(4), 287–300.

Helmreich, R. L., Foushee, C., Benson, R. and Russini, W. (1986). Cockpit resource management: Exploring the attitude-performance linkage. *Aviation, Space, and Environmental Medicine*, **57**, 1198–1200.

Helmreich, R. L., Merritt, A. C. and 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. and Mearns, K. (2006). Safety in shipping: The human element. *Journal of Safety Research*, **37**(4), 401–411.

Holgado Tello, F. P., Chacón Moscoso, S., Barbero García, I. and Sanduvete Chaves, S. (2006). Training satisfaction rating scale. Development of a measurement model using polychoric correlations. *European Journal of Psychological Assessment*, **22**(4), 268–279.

Hong, S. K. and Kim, H. (2016). Bridge resource management training programs in Korea and their effectiveness. *Journal of the Ergonomic Society of Korea*, **35**(4), 237–245.

Hróbjartsson, A., Thomsen, A. S. S., Emanuelsson, F., Tendal, B., Hilden, J., Boutron, I. and Brorson, S. (2013). Observer bias in randomized clinical trials with measurement scale outcomes: A systematic review of trials with both blinded and nonblinded assessors. *Canadian Medical Association Journal*, **185**(4), E201–E211.

Hwang, J. G. and Yang, M. C. (2001). An optimality theory for mid p-values in 2×2 contingency tables. *Statistica Sinica*, **11**, 807–826.

Kent, R. N. and Foster, S. L. (1977). Direct observational procedures: Methodological issues in naturalistic settings. In: Ciminero, A. R., Calhoun, K. S. and Adams, H. E. (eds.). *Handbook of Behavioral Assessment*, New York: John Wiley & Sons, 279–328.

Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training & Development Journal*, **33**(6), 3–9.

Lancaster, H. O. (1961). Significance tests in discrete distributions. *Journal of the American Statistical Association*, **56**(294), 223–234.

Marquardt, N., Robelski, S. and Hoeger, R. (2010). Crew resource management training within the automotive industry: Does it work? *Human Factors*, **52**(2), 308–315.

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E. and 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.

R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Röttger, S., Vetter, S. and Kowalski, J. T. (2013). Ship management attitudes and their relation to behavior and performance. *Human Factors*, **55**(3), 659–671.

Röttger, S., Krey, H., Vetter, S., Stein, M. and Kowalski, J. (2015). What Bridge Teams Do – An Exhaustive Coding Manual for the Assessment of Team Interactions in Seafaring. Poster presented at the Annual Meeting of the HFES Europe Chapter 2015 in Groningen, NL. https://www.researchgate.net/publication/282974408_What_bridge_teams_do_-_an_exhaustive_coding_manual_for_the_assessment_of_team_interactions_in_seafaring

Röttger, S., Vetter, S. and 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.

Saeed, F., Wall, A., Roberts, C., Riahi, R. and Bury, A. (2017). A proposed quantitative methodology for the evaluation of the effectiveness of human element, leadership and management (HELM) training in the UK. *WMU Journal of Maritime Affairs*, **16**(1), 115–138.

Salas, E., Burke, C. S., Bowers, C. A. and Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors*, **43**(4), 641–674.

Salas, E., Benishek, L., Coultas, C., Dietz, A., Grossmann, R., Lazzara, E. and Oglesby, J. (2015). *Team Training Essentials. A Research-Based Guide*. New York: Routledge.

Sanjeev Vakil, K. (2019). Approaches to teamwork and leadership training in maritime education and training institutions: A comparative analysis of the perspectives of seafarers towards teamwork and leadership across different regions. World Maritime University Dissertations, 1176, Malmö, Sweden.

Sellberg, C. and Viktorelius, M. (2020). From technical and non-technical skills to hybrid minds: Reconceptualizing cognition and learning in semi-automated environments. In: Nazir, S., Ahram, T. and Karwowski, W. (eds.). *Advances in Human Factors in Training, Education, and Learning Sciences. Proceedings of the AHFE 2020 Virtual Conference on Human Factors in Training, Education, and Learning Sciences*, Cham, Switzerland: Springer, 191–197.

Staufenbiel, T. (2000). Fragebogen zur Evaluation von universitären Lehrveranstaltungen durch Studierende und Lehrende. [Questionnaire for the evaluation of university courses by students and teachers]. *Diagnostica*, **46**(4), 169–181.

Tvedt, S. D., Espevik, R., Oltedal, H. A., Fjeld, G. P. and Mjelde, F. V. (2018). Can you teach an old seadog new tricks? Experimental evaluation of BRM training in the commercial fleet. *Necesse*, **3**(2), 164–179.

Van Avermaete, J. and Kruijsen, E. (1998). *NOTECHS: The evaluation of non-technical skills of multi-pilot aircrew in relation to the JAR-FCL requirements*. EUROCONTROL NOTECHS Project Final Report: CR, 98443.

Wahl, A. M. and Kongsvik, T. (2018). Crew resource management training in the maritime industry: A literature review. *WMU Journal of Maritime Affairs*, **17**(3), 377–396.

Cite this article: Röttger S, Krey H (2021). Experimental study on the effects of a single simulator-based bridge resource management unit on attitudes, behaviour and performance. *The Journal of Navigation* **74**: 5, 1127–1141. https://doi.org/10.1017/S0373463321000436