

University of Galway Research Repository

An evaluation of the effectiveness of bridge resource management training

Title	An evaluation of the effectiveness of bridge resource management training
Author(s)	O'Connor, Paul
Publication Date	2011
Publication information	O'Connor, P. (2011). "An evaluation of the effectiveness of bridge resource management training". International Journal of Aviation Psychology, 21(4), 357-374.
Publisher	Taylor and Francis
Link to publisher's version	http://www.tandfonline.com/doi/abs/10.1080/10508414.2011.606755
Item record	http://hdl.handle.net/10379/2541

Cite As: O'Connor, P. (2011). An evaluation of the effectiveness of bridge resource management training. *International Journal of Aviation Psychology*, 21(4), 357–374.

**ASSESSING THE EFFECTIVENESS OF BRIDGE RESOURCE MANAGEMENT
TRAINING**

Paul O'Connor

KEY WORDS: Crew resource management, bridge resource management, training evaluation

RUNNING HEADER: Effectiveness of Bridge Resource Management

Corresponding Author: Paul O'Connor, Centre for Innovation and Structural Change, J.E. Cairnes School of Business & Economics. National University of Ireland, Galway, Ireland.
Email: paul.oconnor@nuigalway.ie; poc73@hotmail.com. Tel: +353 (0) 91 492897 Fax: +353 (0) 91 495524.

Accepted for publication in: *International Journal of Aviation Psychology*.

ASSESSING THE EFFECTIVENESS OF BRIDGE RESOURCE MANAGEMENT TRAINING

ABSTRACT

Bridge Resource Management (BRM) is the maritime equivalent of crew resource management (CRM), and has been used in the civilian maritime industry for over a decade. An evaluation was carried out of the effectiveness of the U.S Navy's BRM training by assessing the attitudes towards, and knowledge of, the human factors that contribute to accidents in high-risk organizations. A comparison was made between Surface Warfare Officers (SWOs) who had, and had not, attended BRM training. The responses of the SWOs were also compared to CRM trained U.S Naval aviators. It was found that BRM training did not have a significant effect on the attitudes and knowledge of SWOs. Further, naval aviators were significantly more knowledgeable, and generally held attitudes that were significantly more positive towards the human factors that are causal to accidents in high-risk organizations than SWOs. It was concluded that the Navy's BRM training is not having the impact on knowledge and attitudes that is typically of the CRM training reported in the literature. It is proposed that the main reason for the lack of effectiveness of the BRM program is that the content of the training was not based upon a needs assessment carried out within the surface warfare community.

INTRODUCTION

Crew Resource Management (CRM) training is the most widely applied team training strategy in the world. From the first program in *United Airlines* in 1981 (Helmreich, Merritt, & Wilhelm, 1999), CRM is now being used in a range of high-reliability industries unrelated to aviation (see Hayward & Lowe, 2010 for a review). However, despite the fact that CRM training has been used in U.S. Naval aviation for over two decades, it is only comparatively recently that CRM has been applied in another U.S Navy domain.

Bridge Resource Management (BRM) was introduced into the curriculum of the Surface Warfare Officers School (SWOS; the command that trains officers who will work on naval ships) in the last few years. Although there are a number of papers that have examined the effectiveness of the U.S. Naval aviation program (see O'Connor, Hahn, & Salas, 2010 for a review), there are no reported evaluations of the effectiveness of the U.S Navy BRM program. This paper will briefly describe the U.S Navy's CRM and BRM programs, and report the findings from two studies that assess the effectiveness of the BRM program. The implications of the findings of these studies for the U.S navy and the maritime industry are discussed.

BRM and CRM in the U.S Navy

BRM in the Surface Warfare community. BRM is the maritime equivalent of CRM, and has been used in the civilian maritime industry for over a decade (Hetherington, Flin, & Mearns, 2006). BRM was initially designed to improve the relationship between the Master (ship's Captain) and the Pilot (a mariner who comes aboard to guide a ship through dangerous or congested waters), but soon transformed into the BRM course being taught today which

addresses the broader implications of human factors on the safety and performance of mariners (Barnett, Pekcan, & Gatfield, 2005).

BRM was introduced into the curriculum of Surface Warfare Officers (SWOs) in 2006 (O’Connor, Hahn, & Nulymeyer, 2010). SWOs are U.S. Naval officers whose duties are concerned with the operation of naval ships. The Navy’s BRM curriculum was adapted directly from civilian BRM training. The topics, objectives, and amount of time spent on each topic in the U.S. Navy’s BRM program is shown in Table 1.

Table 1. U.S Navy BRM Course training objectives (Carter-Trahan, 2009).

Topic	Objective	# of Hours
Introduction	Provide student with an overview of BRM.	2
Shiphandling	Provide the student the theory and application ship control.	2
Communication	Demonstrate on open communication style conducive to a comfortable bridge environment.	1
Error Chain	Identify error chains and show how to stop them from causing a mishap.	4
Pilot	Student will know how to execute effective Pilot/CO/OOD relationships.	1
Leadership	Student will know the elements and principles of Leadership.	2
Voyage Planning	Student will understand proper voyage planning.	2 + (2hrs simulator)
Simulator Sessions	Channel transit with moderate environment and traffic.	4
Simulator Sessions	Channel transit with moderate environment and heavy traffic.	8
Simulator Sessions	Execute plan from voyage planning.	6

It is important to indicate that in contrast to the large amount of research that was carried out in the development of the U.S. Navy’s aviation CRM program, *“a review of the literature reveals that there appears to be no empirical foundation for this type of course [BRM] beyond research that was originally conducted in the formation of aviation CRM courses”* (Hetherington

et al, 2004: 407). Therefore, the Navy's CRM program is neither based upon research carried out within the surface warfare community, nor within the maritime industry.

The literature on the implementation of BRM training is far smaller than that on aviation CRM training. There have only been three reports of evaluations of the effectiveness of BRM training (none of which have been published in journals). Byrdorf (1998) reported that over a four year period, BRM training (the topics covered included resource management, assertiveness, communication, team work, and stress coping) lead to a reduction in the mishap rate from one nautical casualty per 30 ship years to one per 90 ship years, and lead to a 15% reduction in insurance premiums.

Fonne and Fredriksen (1995) evaluated a two day BRM training course (no details of topics covered were provided) delivered to 432 high speed marine craft navigators from several Norwegian shipping companies. The reactions to the training were positive, and there was a positive attitude change as measured by a 31 item questionnaire. This change was still evident six months after the course (although moderated).

Brun et al (2001) evaluated a three day BRM training provided to 24 cadets from the Royal Norwegian Naval Academy. The training covered: coping in critical situations, communication and decision making, and team working. It was hoped that the training would increase shared mental models. The reactions to the training were generally positive. However, there was no systematic difference between score on a shared mental models questionnaire pre- and post test in the trained group, or when compared to an untrained group.

CRM in Naval Aviation. The CRM program in U.S. Naval aviation is very mature in comparison to the U.S Navy's BRM program. The goal of naval CRM training is to "*improve mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and*

optimizing risk management” (Chief of Naval Operations, 2001). Every naval aviator must receive ground training and a CRM evaluation during an actual, or simulated flight, by a CRM instructor, or facilitator, once a year. The U.S. Naval aviation CRM training program is based upon research that was carried out in the early 1990s that identified seven critical skill areas to be trained: decision making, assertiveness, mission analysis, communication, leadership, adaptability/ flexibility, and situational awareness (Prince & Salas, 1993).

The U.S. Navy CRM program is centrally controlled, but each of the 48 aviation platform (types of aircraft the Navy and Marine Corps flies) is given the latitude to tailor the CRM program for its particular aircraft and mission. However, the program must adhere to basic naval CRM academic principles, skills and program standardization (O’Connor, Hahn, & Salas, 2010).

The effectiveness of the Navy’s CRM training has been evaluated in the past. O’Connor, Hahn, and Salas (2010) summarize the findings of eight studies that were carried out in the 1990s. Course participants were found to be enthusiastic in their reactions to the training, there was a positive shift in attitudes to the topics addressed in CRM training, there was an increase in knowledge, and improvement in CRM behavior of aircrew as a result of attending the training.

Evaluating CRM/BRM Effectiveness

The measurement of CRM training effectiveness is important to assess whether it is having a positive effect on safety and performance. Kirkpatrick’s (1976) hierarchy is a commonly used model for guiding training evaluation. It consists of four levels.

- Level one (reactions) is concerned with how the participants react to the training. Evaluating reactions is the equivalent of measuring customer satisfaction.

- Level two (learning) is the second level in the hierarchy, and is made up of two components: attitudinal change and knowledge gains. This level is concerned with whether the participants have acquired knowledge, or have modified their attitudes or beliefs as a result of attending the training course.
- Level three (behavior) is the assessment of whether knowledge learned in training actually transfers to behaviors on the job, or a similar simulated environment.
- Level four (organization) is concerned with the organizational level is the highest in Kirkpatrick's (1976) hierarchy. The ultimate aim of any training program is to produce tangible evidence at an organizational level, such as an improvement in safety and productivity.

Although not common in the maritime industry, a large numbers of studies evaluating the effectiveness on CRM training in aviation (military and civilian), healthcare, and other domains have been carried out (see O'Connor, Flin, & Fletcher, 2002; Salas, Fowlkes, Stout, Milanovich, & Prince, 1999; Salas, Wilson, Burke, & Wightman, 2006 for reviews). Nullmeyer, Spiker, Wilson, and Deen (2003) assert that CRM trainers struggle to maintain the resources necessary to support CRM training. So, if evidence of effectiveness is found, this will be a useful metric for provide evidence in support of continued, or increased, resourcing of the program. Further, by identifying strengths and weaknesses the program, this provides evidence for where improvements need to be made. The Federal Aviation Authority (FAA) states that for CRM training *"it is vital that each training program be assessed to determine if CRM training is achieving its goals. Each organization should have a systematic assessment process. Assessment*

should track the effects of the training program so that critical topics for recurrent training may be identified and continuous improvements may be made in all other respects” (FAA, 2004: 12).

Study Purpose

The purpose of this study was to examine the effectiveness of the Navy’s BRM training program. This was achieved by carrying out three different comparisons.

1. The knowledge and attitudes (learning level) of SWOs who have received BRM training was compared to SWOs that have not received BRM training.
2. The knowledge and attitudes of SWOs who had not received BRM training was compared to CRM trained aviators.
3. The knowledge and attitudes of CRM trained naval aviators was compared with BRM trained SWOs.

Study limitations

There are a number of limitations to this study. Firstly, the lack of data from a group of naval aviators who had not received CRM training limits the conclusions that can be drawn from the comparison between the two communities. It is possible that the naval aviators may have more knowledge of, and more positive attitudes towards, CRM prior to training than SWOs. Every naval aviator is mandated to receive CRM training. From a scientific perspective, it would have been desirable to withhold CRM training from a number of aviators. Even in the extremely unlikely event of senior leadership agreeing to this proposal, given the findings supporting the effectiveness of CRM training reported in the literature, it would not be ethical.

A related limitation is that as can be seen from the description of the BRM and CRM the content of the training is not the same. Therefore, it could be argued that the between groups comparison is not comparing 'like with like'. This argument is certainly true. However, although the training is different, the purpose is the same- to improve safety and performance. If one of the training programs is shown to have a greater effect on knowledge and attitudes than the other, this suggests that the training with the lower effect size is worthy of scrutiny. Also, other reviews of the CRM evaluation literature (O'Connor et al., 2002; O'Connor et al, 2008; Salas, et al, 1999; Salas, et al, 2006) have treated the CRM training across the studies reviewed as equal (although it is certainly not) in order to draw conclusions about the effectiveness of the training.

Finally, it is also recognized that it would have been desirable to carry out evaluations at more than one level of Kirkpatrick's (1976) training hierarchy. However, the reactions of participants' are not systematically assessed, and evaluations of behaviors are not consistently carried out in either the surface warfare or aviation communities. Initially it was thought that it would be possible to compare the human factors mishap rate between the aviation and surface warfare communities. Considerable detail is collected on the causes of U.S Naval aviation mishaps (although recently concerns have been raised about the reliability of the mishap coding system; O'Connor & Walker, 2011) the same is not true in the surface warfare community. Carter-Trahan (2009) examined 111 major mishaps in the surface warfare community investigated by the Naval Safety Center from 1999 to 2009. Only 23% of mishaps were attributed to human factors (the remaining were attributed to material causes 12%, and unknown causes 65%). Given the large proportion of mishaps attributed to "unknown" causes, and a human factors mishap rate that is far lower than would be expected, the reliability of the human factors analysis of the surface warfare mishaps would appear to be questionable. So, the

evaluation reported in this study was limited to the learning level of Kirkpatrick's (1976) training evaluation hierarchy.

These limitations are an unfortunate issue that is associated with carrying out research in an applied domain. Applied researchers often do not have the luxury of setting up a study as scientifically as he/she would like. Nevertheless, despite the limitations, given the lack of evaluations of BRM reported in the literature it is proposed that the two evaluation studies described in this paper provide insight into the effectiveness of the U.S. Navy's BRM program and has implications for the effectiveness of BRM training more generally.

ASSESSMENT OF KNOWLEDGE OF CRM

Knowledge: Introduction

Knowledge assessment is a useful method for evaluating the effects of training (Flin et al., 2008). In a meta-analysis of four CRM studies reporting a knowledge evaluation, a mean Cohen's d of 0.59 was found (O'Connor et al, 2008). Cohen (1988) hesitantly defined effect sizes as small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$). The current study also used a multiple-choice knowledge test to assess the CRM knowledge of aviators and SWOs.

Knowledge: Method

Test development. A ten item multiple-choice test was developed to address human factors issues that have been identified as causal to accidents in both the aviation and maritime environments (situation awareness, decision making, communication, stress, and fatigue). The items were designed to be generic so they would make sense to both aviators and SWOs. The questionnaire was administered as a pilot test on ten naval aviators, and ten SWOs.

Procedure. The questionnaire was distributed to naval aviators who were students at the Naval Aviation Schools Command (NASC), Pensacola, Florida and the Naval Postgraduate School (NPS), Monterey, California and qualified SWOs attending training at SWOS, Newport, Rhode Island, and NPS. Information on the time since CRM/BRM training was not collected. However, the questionnaire was not distributed immediately after CRM or BRM training, so can be considered the baseline knowledge of participants. The study protocol was approved in advance by the Institutional Review Board (IRB) of NPS.

Participants. A total of 142 responses were obtained from SWOs who had attended BRM training, and 24 from SWOs who indicated that they had not received BRM training. A total of 183 responses were obtained from naval aviators (all of whom had received CRM training). All of the respondents were junior officers.

Knowledge: Results

Naval aviators got a mean of 68.7% (sd= 14.5) questions correct, compared to a mean of 55.0% (sd= 11.9) correct for BRM trained SWOs, and 50% (sd= 10.8) for SWOs who had not attended BRM. A one-way between subjects Analysis of Variance (ANOVA) was used to assess whether there was a significant differences in knowledge scores for aviators, BRM trained SWOs, and non-BRM trained SWOs. Due to concerns about homogeneity of variance and unequal sample size, Welch's (1951) correction to F was used, resulting in a significant finding ($F(2,68)= 51.5, p<.05$). As would be expected, post-hoc analysis using Tamhane's T2 indicated that the aviators scored significantly better than both groups of SWOs. The difference between BRM and non-BRM trained SWOs was not significant. The effect size of the difference between the aviation and BRM and non-BRM trained SWOs was a Cohen's *d* of 1.02 and 1.42

respectively- a large effect size. The effect size of the difference between the SWOs that had received BRM training, and those that had not was a Cohen's d of 0.44 (a small to medium effect size).

Knowledge: Discussion

Knowledge is an important part of learning. It is acknowledged that a multiple-choice test is a fairly coarse assessment of knowledge. However, it is consistent with how knowledge has been assessed in other CRM evaluation studies (e.g. Salas, et al, 1999), and provides a broad measure of the how much the participants knew about common human factors causes of accidents in high-risk organizations.

Differences in knowledge between SWOs who had attended BRM training and those that had not. There was not a significant difference on the knowledge test based upon whether the SWOs had attended BRM training. This is not entirely unsurprising given that, as can be seen from the BRM training objectives outlined in Table 1, the majority of the topics addressed by the knowledge test were not explicitly addressed in BRM training. The failure to address all of these topics would seem to be a large weakness of the BRM training given that research conducted in the maritime industry (see Hethrington et al, 2006) as well as other high-risk organizations (Flin, O'Connor, Crichton, 2008) would suggest that loss of situation awareness, poor decision making, poor communication, and the effects of stress and fatigue are major contributors to accidents.

Differences in knowledge between naval aviators and SWOs. There was a large effect size in the difference between the knowledge of aviators and SWOs who had, and had not, received BRM training. This effect size was greater than that medium effect size of knowledge reported in O'Connor et al.'s (2007) meta-analysis of CRM training effectiveness.

The results of the knowledge assessment are encouraging for the aviation community because despite the fact that CRM knowledge is not something that is routinely tested, aviators seem to retain the knowledge they are taught in squadron CRM training. Nevertheless, the performance is lower than would be desired. Although not currently carried out in naval aviation, it is suggested that just as is the case for knowledge of other aviation related information (e.g. knowledge of systems and emergency procedures), CRM knowledge should be tested.

ASSESSMENT OF CRM ATTITUDES

Attitudes: Introduction

The most commonly used tools for assessing pilot attitudes towards the skills taught in CRM training are the Cockpit Management Attitudes Questionnaire (CMAQ; Gregorich & Wilhelm, 1990), and the Flight Management Attitude Questionnaire (FMAQ; Helmreich & Merritt, 1998). These surveys are designed to assess attitudes toward stress, hierarchy, teamworking, and human error. Previous research has found that the CMAQ and FMAQ are relevant to understanding error, predictive of human performance, and sensitive to training interventions (Sexton, Thomas, & Helmreich, 2000). However, apart from a few isolated examples (e.g. Gregorich et al 1990; Taylor, 2000), an examination of the psychometric properties of these questionnaires is rarely reported. Particularly when being applied to a new domain, it is important to examine the psychometric properties of a measurement instrument.

A meta-analysis of nine studies in which an attitude assessment was used to evaluate CRM training, showed a large effect size (O'Connor et al, 2008). The current study will compare the attitudes of naval aviators and SWOs. The psychometric properties of the instrument that was

used will be examined, and an attempt made to identify a measurement model that is invariant across both populations.

Attitudes: Method

Instrument development. A 31 item Navy/Marine Corps aviation version of the FMAQ questionnaire was developed for this study called the Naval Aviator Human Factors (NAHF) questionnaire (see O'Connor & Jones, 2009 for more details). For each statement, the degree to which participants agree was assessed with a 5-point Likert scale. Following a preliminary assessment of the psychometric properties of 364 responses by U.S Naval aviators to the questionnaire, nine items were dropped for analysis based upon high levels of skewness and kurtosis (see O'Connor & Jones, 2009). The remaining 22 items in the revised NAHF questionnaire consisted of four proposed factors:

- My stress- emphasizes the consideration of- and possible compensation for- stressors in oneself.
- Stress of others- emphasizes the consideration of- and possible compensation for- stressors in other team members
- Communication- encompasses communication of intent and plans, delegation of tasks and assignment of responsibilities, and the monitoring of team members.
- Command responsibility- appropriate leadership and its implications for the delegation of tasks and responsibilities.

The 22 item revised NAHF questionnaire was adapted for the SWO community by the author and an experienced SWO. The revised questionnaire was then piloted on 10 SWOs.

Procedure. The NAHF questionnaire was distributed to U.S. Naval aviation at either the NASC or NPS. The surface warfare version of the NAHF questionnaire was distributed to SWOs at SWOS and at NPS. The study protocol was approved in advance by the Institutional Review Board (IRB) of NPS.

Participants. Responses were obtained from 526 U.S. Naval aviators (all of whom had attended CRM training). Examining the distribution of respondents by rank, 38% were junior officers, and 62% senior officers. A total of 142 responses were obtained from SWOs who had attended BRM training, and 24 SWOs who indicated that they had not received BRM training. All of the SWOs were junior officers.

Psychometric Analysis. A CFA technique as implemented by EQS for windows was used to establish the measurement invariance of the questionnaire across the two populations. This analysis was carried out using the three hierarchical steps described by Byrne (1994). The first step is to establish the baseline models. This process is carried out independently for each group. It is an assessment of whether the factors, and the items that load on them, are the same in both samples. The adequacy of the factor model is assessed through a number of fit indices (see Byrne, 2006, for a discussion). Byrne (2006) states that although the models should be similar, they do not have to be identical. The initial step for testing invariance requires only that the same number of factors load on the same pattern of variable in each group. In the next step the configural invariance is assessed by testing a multigroup representation of the baseline models. In this process invariance tests are conducted across the two groups simultaneously, and provides a baseline value against which subsequent models are compared (Byrne, 2006). The purpose of this step is to ensure there are equal numbers of factors and the same items load on these factors in both groups. Finally, the measurement invariance is assessed. This step ensures the

relationship between items and factors is the same across the two groups. The primary interest is on the factor loadings.

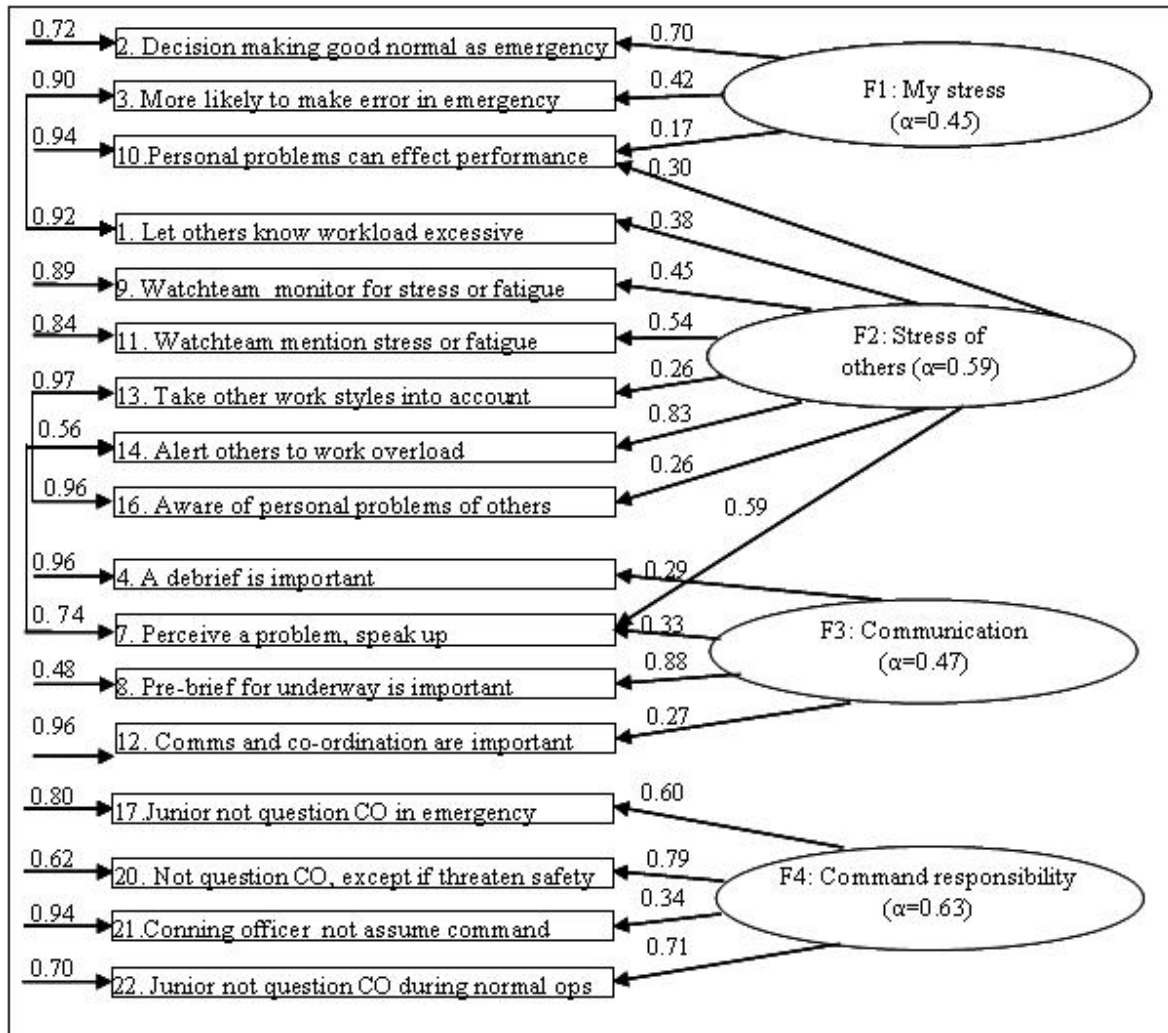
Between group comparisons. Consideration was given to using the one of the refined methods for calculating the factor scores provided in EQS. However, it was decided to simply take the average of the item for each factor. The reason for choosing this method for calculating the factor scores was this was what was done in the previous studies of the CRM attitudes of naval aviators, and non-refined factor scores are thought to be more stable across samples than refined methods (Grice & Harris, 1998). Also, as O'Connor and Jones (2009) reported significant differences on NAHF questionnaire responses based upon the rank of aviators and the sample of SWO did not contain any senior officers, it was decided that only the 199 responses by junior officer in the naval aviation sample would be used for the between groups comparison.

Attitudes: Results

Psychometric Analysis. Firstly, a baseline model was established for the aviation sample using the 22-item revised NAHF. However, the fit of the model was not found to be acceptable for the SWO sample ($\chi^2 = 538$, $df = 190$, $p > 0.05$; CFI(robust) = 0.78; GFI = 0.85; and RMSEA(robust) = 0.051). After examining the results of the Wald test (used to identify items that should be dropped from the model), and talking to SWOs about the items, it was decided to drop items 18 (*The Officer of the Deck should take physical control and drive the ship in emergency and non-standard situations*) and 19 (*the Officer of the Deck should verbalize plans for procedures or maneuvers and should be sure that the information is understood and acknowledged by other crew members*). The SWOs felt that these statements did not make sense as applied to the surface warfare community.

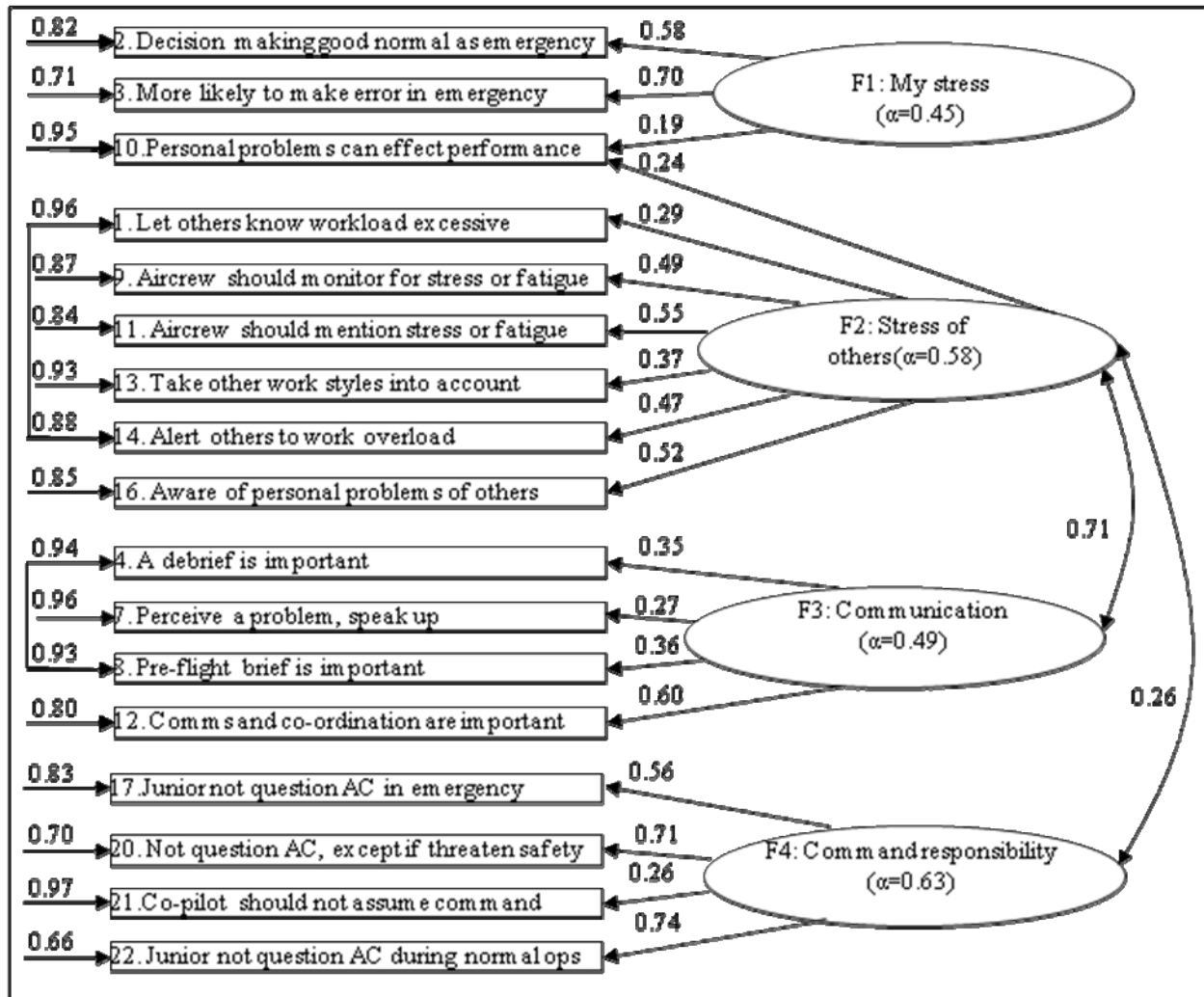
The model was then re-run, and after removing item 15 (*the specific roles and responsibilities of the watchteam in an emergency are identified during the preflight brief*), allowing item 7 (*if I perceive a problem during underway operations, I will speak up, regardless of who might be affected*) to cross load onto the ‘stress of others’ factor, and allowing covariance between three pairs of error terms a reasonable fit was obtained ($\chi^2 = 461$, $df = 136$, $p > 0.05$; CFI(robust) = 0.92; GFI = 0.88; and RMSEA(robust) = 0.035). This model is shown in Figure 1. It should be indicated that although the GFI value is not greater than 0.9 (indicative of an acceptable fit), the fit on the other indices was acceptable. Therefore, it was decided to use this new model as the revised baseline model for the aviation population.

Figure 1. SWO sample standardized solution, and Cronbach’s Alphas.



The initial fit of the revised baseline model for the aviation data was not acceptable ($\chi^2 = 1161$, $df = 136$, $p > 0.05$; CFI = 0.85; GFI = 0.95; and RMSEA = 0.005). However, by allowing covariances between the ‘stress of others’ factor and the ‘communication’ and ‘command responsibility’ factor and between two pairs of error terms an acceptable fit was achieved ($\chi^2 = 1623$, $df = 272$, $p > 0.05$; CFI = 0.93; GFI = 0.95; and RMSEA = 0.037; see figure 2).

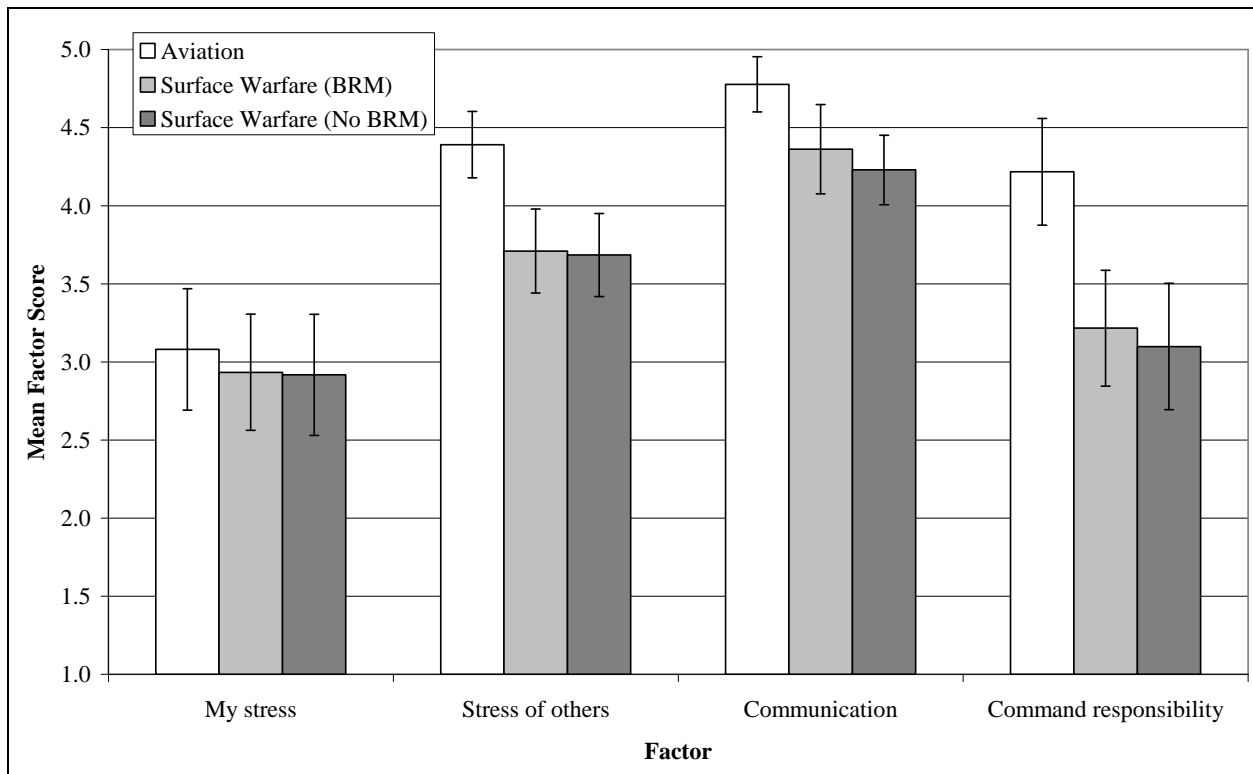
Figure 2. Aviation sample standardized solution, and Cronbach's Alphas.



The baseline models were testing for configural invariance, and an acceptable goodness of fit was found ($\chi^2 = 1161$, $df = 136$, $p > 0.05$; CFI(robust) = 0.94; GFI = 0.96; and RMSEA(robust) = 0.027). Finally, the models were tested for measurement invariance, and an acceptable fit was found ($\chi^2 = 1623$, $df = 272$, $p > 0.05$; CFI(robust) = 0.91; GFI = 0.94; and RMSEA(robust) = 0.032).

Between group comparisons. The mean and standard deviations for each factor score are shown in Figure 3. Given that item 7 cross loaded on two factors in the SWO model, but did not cross load in the aviation model, the decision was made not to include it in the calculation of the ‘stress of others’ factor.

Figure 3. Mean and standard deviation of attitude scale scores for aviators and SWOs



Between subjects ANOVAs were used to assess whether there were significant differences in factor scores for aviators, BRM trained SWOs, and SWOs who had not received BRM training. Due to concerns about homogeneity of variance and unequal sample size, Welch’s (1951) correction to F was used. The results of the ANOVAs and the effect sizes (calculated using Cohen’s *d*) are shown in Table 2.

Table 2. Welsh's F values and effect sizes for differences between groups for factor scores.

Factor	Welsh's F	Cohen's <i>d</i>		
		Aviation-BRM trained SWOs	Aviation-no BRM SWOs	BRM- no BRM SWOs
My stress	F(2,57)= 2.07, n.s.	0.18	0.41	0.03
Stress of others	F(2,55)= 95.7, p<.05	1.41	1.48	0.04
Communications	F(2,54)= 40.1, p<.05	0.87	1.38	0.29
Command responsibility	F(2,57)= 110.1, p<.05	1.42	1.48	0.14

For all of the significant ANOVAs post-hoc analysis was carried out using Tamhane's T2. It was found that where there were significant differences in the factor scores, these were due to the difference between the aviators and both groups of SWOs. The differences between the SWOs who had attended BRM and those that had not were not significant.

Attitude: Discussion

Psychometric analysis. Using post-hoc analysis, and discussions with SWOs, a reasonably fitting CFA model was found for the SWO data. A broadly similar model was also found to fit the naval aviation data. However, as can be seen from figures 1 and 2, the models that were fit to the aviation and SWO data sets were not exactly the same. Although the requirement for the same error covariances is regarded as being excessively stringent (Byrne, 2006), it can be seen that the cross loadings of item 7 on the SWO model, and the correlations between the factors in the aviation model represent more substantial differences between the models. Therefore, there is only partial-measurement invariance between the two models. The extent to which this invalidates a claim of measurement invariance is of some debate in the technical literature (Byrne, 2006).

Regardless of the partial-measurement invariance debate, it is suggested that there is a need for further refinement to the attitude questionnaire. The sample size of the SWO group was at the low end for this type of analysis. A larger sample should be collected (to include senior officers), and the model fit reassessed. Also, the reliability of the factors, although typical of this type of questionnaire (for example, Gregorich et al, 1990, reported alphas of between 0.47 and 0.67 for the CMAQ), is lower than would be desired.

Differences in attitudes between SWOs who had attended BRM training and those that had not. No significant differences in the factor scores were found between SWOs who had received BRM training and those who had not. Similar to the lack of a significant difference between the same two groups on the knowledge test, it is proposed that this finding can be attributed to the limited scope of the U.S Navy's BRM training in addressing topics that have been shown to be causal to accidents in high-risk domains.

Differences in attitudes between naval aviators and SWOs. For all of the factors except for 'my stress', naval aviators had attitudes that were significantly more sensitive to the concepts addressed in CRM training than SWOs- both those that had received BRM training, and those that had not. Further, where the differences were significant, the effect size of the difference between naval aviators and SWOs was higher than is generally reported in other CRM evaluation studies (see O'Connor et al., 2008).

The differences in the mean factor scores for 'my stress' and 'stress of others' would appear to suggest that neither populations wished to acknowledge the effect of stressors on their own performance, although they are willing to regard stressors as effecting the performance of others. At least in the naval aviation population, the low mean factor score on the 'my stress' as compared to the other factors is not unexpected. When faced with a stressful situation, military

aviators tend to use avoidance stress coping strategies such as denial, distraction, repression, and suppression (see Campbell & O'Connor, 2010 for a detailed discussion). There is no research to identify whether avoidance coping is utilized in the SWO community beyond anecdotal evidence. However, the stress coping strategies utilized by SWOs may be worth examining so that the BRM course can be tailored with this in mind.

The difference between naval aviators and SWOs was significant for the 'communication' factor. Nevertheless, it is encouraging that both communities recognized the importance of communication as indicated by the high proportion of favorable responses to these items. The failure to exchange information and co-ordinate actions is one factor that differentiates between good and bad team performance (Driskell & Salas, 1992).

The large difference in the effect size for the 'command responsibility' factor could be attributed to the difference between the culture, structure, and operating environment of the two communities. It is common in naval aviation to say that 'there is no such thing as rank in the cockpit'. In reality, this is not always the case. However, this is certainly never the case in the SWO community. A ship's crew has many more people and is much more hierarchical than a small aircraft crew. Nevertheless, it is suggested that the need for assertiveness by junior officers, and listening by senior officers is something that may be worthy of increased focus in the BRM program.

GENERAL DISCUSSION

Although there are limitations to the evaluation studies reported in this paper, it can be concluded that the Navy's BRM training is not having the impact on knowledge and attitudes that is typically of the CRM training reported in the literature. It is proposed that the main reason

for the lack of effectiveness of the BRM program is because the content of the training was not based upon a needs assessment carried out within the surface warfare community. In fact, as stated in the introduction, the U.S Navy's BRM program is not even based upon a needs assessment conducted in the maritime industry, but is established upon research carried out in civilian aviation. A needs assessment is the first stage of any model of training development (e.g. Goldstein & Ford, 2002; Truelove, 1997). As stated by Helmreich, et al (1999): "*I am not suggesting the mindless import of existing programs; rather, aviation experience should be used as a template for developing data driven actions reflecting the unique situation of each organization* (p784)". However, similar to the situation when CRM training was first introduced in military and civil aviation, there is a lack of human factors research conducted in the maritime domain to inform the content of BRM training.

In their review of the maritime human factors literature Harrington et al (2006) only found 20 relevant studies and stated that "*there are many "gaps" in the maritime literature, and a number of methodological problems with the studies undertaken to date*" (p. 410). Schröder-Hinrichs (2010) also noted that it was noticeable that recently human factors had not been high on the international maritime agenda. Therefore, more human factors research is required within the maritime industry, and the surface warfare community to identify the topics that need to be trained in BRM, rather than basing the content on research carried out in the civil aviation industry.

Another important part of any model of training development is the need to evaluate training effectiveness. As discussed in the introduction, compared to aviation, there are very few studies reported in the literature in which an evaluation of the effectiveness of BRM training has been reported. In the absence of training evaluation, whether the training is positively impacting

the safety or performance of operators is unknown. There is a great deal of information available on how CRM training can be evaluated (e.g. Flin, et al, 2008). This knowledge should be used to assess the impact of CRM training at multiple levels of Kirkpatrick's (1976) evaluation hierarchy.

CONCLUSIONS

The temptation to adapt a CRM program developed in one domain and apply it in another must be resisted by developers of CRM training. This practice risks the development of ineffective training programs that fail to impact operator performance. The design and delivery of training is a science (Salas & Cannon-Bowers, 2000). Unlike when CRM training was first applied in civil aviation, there are now solid principles, guidelines, tips and specifications that can be used to develop CRM programs (Salas et al, 2006). Researchers must use this knowledge to design CRM training programs that have a positive effect on the safety and performance of operational personnel.

ACKNOWLEDGEMENTS

The information presented is the opinion of the author and does not necessarily reflect that of the U.S. Navy, or the National University of Ireland, Galway.

REFERENCES

- Barnett, M. L., Pekcan, C. H., & Gatfield, D. I. (2005). Content and Context: Understanding the Complexities of Human Behaviour in Ship Operation. *Proceeding of the Human Factors in Ship Operations Conference* (pp. 1-11). London: The Royal Institution of Naval Architects.
- Brun, W., Eid, J., Johnsen, B.H., Ekornås, B., Laberg, J.C. & Kobbeltvedt, T. (2000). Shared mental models and task performance: Studying the effects of a crew and bridge resource management training program. (Project report: 1 2001). Amsterdam, Norway: Militaer Psykologi og Ledelse.
- 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*. Paper presented at the 23rd Conference of the European Association for Aviation Psychology, Vienna.
- Byrne, B.M. (2006). *Structural Equation Modeling with EQS*. Mahwah, NJ: Lawrence Erlbaum.
- Campbell, J. & O'Connor, P. (2010) Coping with stress in military aviation: A review of the research. In: P. O'Connor & J. Cohn (Eds.). *Human Performance Enhancements in High-Risk Environments: Insights Developments, and Future Directions from Military Research* (pp.169-188). Santa Barbara, CA: ABC- Clio.
- Carter-Trahan, A. (2009). *An Examination of the Human Factors Attitudes and Knowledge of Surface Warfare Officers*. Masters thesis. Monterey, CA: Naval Postgraduate School.
- Chief of Naval Operations (2001). *Crew Resource Management Program, OPNAVINST 1542.7C*. Washington DC: Author.

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Driskell, J.E., Salas, E. (1992). Collective behavior and team performance. *Human Factors*, 34, 277-288.
- Federal Aviation Administration (2004). *Advisory Circular No 120-51E: Crew Resource Management Training*. Washington, DC: Author.
- Flin, R., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: training non-technical skills*. Aldershot, England: Ashgate Publishing Ltd.
- Fonne, V. M. & Fredriksen, O. K. (1995). Resource management and crew training for HSV-navigators. In R. Jensen (Ed.) *Proceedings of the 8th International Symposium of Aviation Psychology*, Columbus, Ohio, 585-590.
- Goldstein, I.L. & Ford, K.J. (2002) *Training in organizations: needs assessment, development, and evaluation*. Belmont, CA: Wadsworth.
- Gregorich, S. E., Helmreich, R. L., & Wilhelm, J. A. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology*, 75, 682–690.
- Grice, J. W. & Harris, R. J. (1998). A comparison of regression and loading weights for the computation of factor scores. *Multivariate Behavioral Research*, 33(2), 221-247.
- Hayward, B.J. & Lowe, A.R. (2010). The migration of crew resource management training. In: R.L. Helmreich, B.G. Kanki, & J. Anca (Eds.). *Crew Resource Management*. (pp. 317-344). San Diego. CA: Elsevier.
- Helmreich, R. L., & Merritt, A. C. (1998). *Culture at Work in Aviation and Medicine: National, Organizational and Professional Influences*. Aldershot: Ashgate.
- Helmreich, R., Merritt, A., & Wilhelm, J. (1999). The evolution of crew resource management

- training in commercial aviation. *International Journal of Aviation Psychology*, (1), 19-32.
- Hetherington, C., Flin, R., & Mearns, K. (2006). Safety in shipping: The human element. *Journal of Safety Research*, 37, 401-411.
- Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig & L. R. Bittel (Eds.), *Training and development handbook* (pp. 18.1-18.27). New York: McGraw Hill.
- Nullmeyer, R., Spiker, V.A., Wilson, D., & Deen, G. (2003). Key crew resource management behaviors underlying C-130 aircrew performance. In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference* [CD-ROM]. Arlington VA: National Training Systems Association.
- 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. *International Journal of Aviation Psychology*, 18(4), 353-368.
- O'Connor, P., Flin, R. & Fletcher, G. (2002) Methods used to evaluate the effectiveness of CRM training: A literature review. *Journal of Human Factors and Aerospace Safety*, 2, 217-234.
- O'Connor, P., Hahn, R., & Nullemeier, R. (2010). The military perspective. In: R.L. Helmreich, B.G. Kanki, & J. Anca (Eds.). *Crew Resource Management*. (pp. 445-468). San Diego, CA: Elsevier.
- O'Connor, P., Hahn, R. & Salas, E. (2010) The U.S. Navy's Crew Resource Management program: the past, present, and recommendations for the future. In: P. O'Connor & J. Cohn (Eds.). *Human Performance Enhancements in High-Risk Environments: Insights, Developments, and Future Directions from Military Research* (pp. 90-105). Santa Barbara, CA: ABC- Clio.

- O'Connor, P. & Jones, D. (2009). *The crew resource management attitudes of U.S. Naval aviators*. Paper presented at the Human Factors and Ergonomics Society conference. San Antonio, TX.
- O'Connor, P. & Walker, P. (2011). Evaluation of a human factors analysis and classification system as used by simulated mishap boards. *Aviation, Space and Environmental Medicine*, 82, 44-48.
- Prince, C., & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. Wiener, B. Kanki & R. Helmreich (Eds.), *Cockpit Resource Management* (pp. 337-366). San Diego: Academic Press.
- Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors*, 41, 161-172.
- Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In L. Tobias & D. Fletcher (Eds.), *Training and retraining: A Handbook for business, industry, Government, and the military* (pp. 312-335). New York: MacMillan.
- Salas, E., Fowlkes, J. E., Stout, R. J., Milanovich, D. M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors*, 41(2), 326-343.
- Salas, E., Wilson, K.A, Burke, C.S., & Wightman, D.C. (2006). Does CRM training work? An update, extension and some critical needs. *Human Factors*, 14, 392-412.
- Schröder-Hinrichs, J.-U. (2010). Human and organizational factors in the maritime world – are we keeping up to speed? *WMU Journal of Maritime Affairs*, 9(1), 1-3.
- Sexton, J.B., Thomas, E.J., & Helmreich, R.L. (2000). Error, stress, and teamwork in medicine and aviation: cross sectional surveys. *British Medical Journal*, 320, 745–749

- Taylor, J. C. (2000). Reliability and validity of the Maintenance Resource Management, Technical Operations Questionnaire (MRM/TOQ). *International Journal of Industrial Ergonomics*, 26(2), 217–230.
- Truelove, S. (1997). *Training in Practice*. Oxford: Blackwell.
- Welch, B.L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38, 330-336.