

Review

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

Jop Havinga ^{1,*} , Robert Jan de Boer ² , Andrew Rae ¹  and Sidney W. A. Dekker ¹

¹ Safety Science Innovation Lab, School of Humanities, Griffith University, Nathan 4111, Australia; d.rae@griffith.edu.au (A.R.); s.dekker@griffith.edu.au (S.W.A.D.)

² Aviation Academy, Amsterdam University of Applied Sciences, Amsterdam 1097 DZ, The Netherlands; rj.de.boer@hva.nl

* Correspondence: j.havinga@griffith.edu.au; Tel.: +61-7-3735-4757

Received: 15 June 2017; Accepted: 26 October 2017; Published: 31 October 2017

Abstract: Crew resource management (CRM) training for flight crews is widespread and has been credited with improving aviation safety. As other industries have adopted CRM, they have interpreted CRM in different ways. We sought to understand how industries have adopted CRM, regarding its conceptualisation and evaluation. For this, we conducted a systematic review of CRM studies in the Maritime, Nuclear Power, Oil and Gas, and Air Traffic Control industries. We searched three electronic databases (Web of Science, Science Direct, Scopus) and CRM reviews for papers. We analysed these papers on their goals, scope, levers of change, and evaluation. To synthesise, we compared the analysis results across industries. We found that most CRM programs have the broad goals of improving safety and efficiency. However, there are differences in the scope and levers of change between programs, both within and between industries. Most evaluative studies suffer from methodological weaknesses, and the evaluation does not align with how studies conceptualise CRM. These results challenge the assumption that there is a clear link between CRM training and enhanced safety in the analysed industries. Future CRM research needs to provide a clear conceptualisation—how CRM is expected to improve safety—and select evaluation measures consistent with this.

Keywords: crew resource management (CRM); training; systematic review; maritime; air traffic control; nuclear power industry; oil and gas

1. Introduction

The development of crew resource management (CRM) for flight crews has been repeatedly called a great success story [1,2] and sometimes even “one of the greatest successes of aviation and the human factors/ergonomics field” [3]. While it initially was met with some resistance, CRM is now well accepted and has spread to all countries and most airlines [2]. Given that airlines are often regarded as having exemplary safety records, it is not surprising that other industries have started to adopt CRM training as well [3].

However, evaluative research on different CRM programs has not often found CRM programs successful. Reviews on the subject have not found a clear link between CRM and a reduction of accidents [4]. A key problem is that evaluative studies on CRM suffer from a variety of methodological issues [5], including the selection of outcome measures. When other industries adopt CRM, this makes it hard to recognise successful CRM programs, to understand which parts of a CRM program are important for success, or even to interpret what “success” looks like outside of encouraging specific cockpit behaviours.

This leads to our main research question:

- How has CRM been adopted in the maritime (MAR), the nuclear power (NPI), oil and gas (O&G), and air traffic control (ATC) industries?

To understand the adoption of CRM in an industry, it is first important to understand what the adopters understand CRM to be, as different people can conceptualise CRM training differently. Changes in conceptualisation or definition during the adoption of CRM should then be reflected in the evaluation of CRM programs [6]. Therefore, we have split the main research question into:

- How has CRM been conceptualised in the maritime, the nuclear power, oil and gas, and air traffic control industries?
- How has CRM been evaluated in the maritime, the nuclear power, oil and gas, and air traffic control industries?

2. Methods

The purpose of this systematic literature review is to explore how industries have adopted CRM. Where applicable, we follow the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement [7] on the reporting of systematic reviews. Since our goal was not to assess the efficacy of CRM as a thing-in-itself, but rather to study how CRM has been adopted and evaluated, the elements of the PRISMA guidelines focussed on reaching conclusions about efficacy did not apply. In addition, to lower the cognitive burden on the reader, we describe the review of the conceptualisation and evaluation sequentially, and draw conclusions after each section, before combining the results.

In this section, we will describe our study in terms of (1) the search for articles; (2) the selection process; (3) the review process of the conceptualisation of CRM; and (4) the review process of the evaluation of CRM programs.

2.1. Search Process

To orientate, we familiarised ourselves with the reviews of CRM and textbooks on CRM [2,4,8,9]. From the orientation, we established the initial search terms and explored the domains that were available. We chose to analyse the following fields: the maritime industry (MAR), the NPI, the O&G, and ATC. Industries such as fire-fighting and railway were left out because of the low number of publications. We also excluded military CRM and healthcare CRM because these industries have already been thoroughly and recently reviewed [2,10–14].

We searched for relevant articles using the references of the orientation literature, as well as the search engines Web of Science, Science Direct, and Scopus. The initial search terms came from the orientation literature, but the search terms were updated as new terms were found in the literature. The search included articles up to the end of 2015. The final search terms per industry were:

MAR: (“crew resource management” and maritime) or (“bridge resource management”) or (“bridge team management”) or (“maritime resource management”) or (“engine room resource management”)

NPI: (“crew resource management” and “nuclear power”)

O&G: (“crew resource management” and oil and gas) or (“crew resource management” and offshore) or (“deep water well control”) or (“crew resource management” and “well operations”)

ATC: (“crew resource management” air traffic control) or (“team resource management”) or (“controller resource management”) or (“air traffic team enhancement”) OR (“controller awareness and resource training”) or (ATCRM)

2.2. Selection Process

For our review we used the following criteria for an article to be relevant:

- Written in English;
- Focuses on one of the chosen industries;
- Describes an implemented CRM program or specifies what CRM training should cover. Self-proclaimed partial courses were included, as no definition was found of what makes a full CRM course;
- Makes first hand comments about CRM; studies that combined and reinterpreted other literature were included, but studies that only repeated other CRM literature were excluded.

To filter out all papers not matching our criteria, we used the following process. First, we combined and structured all search results and references found in reviews into a spreadsheet. In this process, duplicates in the search results were taken out. Next, all articles were scanned by reading their abstract to judge whether they could match the selection criteria. In cases of doubt, the article was searched with the search terms of that industry to see how and why CRM was mentioned in those papers. The papers that were left were read in full and assessed for their eligibility. This process was repeated for each of the industries.

Inclusion criteria were agreed by all authors, applied by the first author, and reviewed by the third author.

2.3. Conceptualisation Review

2.3.1. Analysis

We analysed all papers to see how CRM was conceptualised. It is impossible to capture all possible differences between CRM applications, so we determined initial categories using statements made by well-referenced authors, regarding the goals of CRM, the levers of change, and the scope of CRM.

Goals

First, we looked at the major goals of CRM training. For this we went to Lauber's [15] definition of CRM as "using all available resources—information, equipment, and people—to achieve safe and efficient flight operations". This is the oldest definition of CRM, and variations of this definition continue to be used. The 'using all available resources' part is open to multiple interpretations, but 'safe' and 'efficient' provided two clear starting categories. We analysed all papers based on whether they mentioned safety and efficiency as goals. 'Safety' included the reduction of accidents and incidents. As efficiency might not be easily applied to all industries, we translated this to any increase in performance. Statements regarding desired, expected, and found effects were all counted as indicators of goals.

Levers of Change

Next, we investigated the levers of change—the process through which CRM was expected to achieve its larger goals. Based on the orientation literature, we were interested in two particular attributes of the process:

- The unit of analysis for the immediate effects of CRM [16], and
- The stance of CRM with respect to compliance.

Flin et al. [16] highlight the unit of analysis (e.g., individual or team) as a contested issue in the conceptualisation of CRM. Changes in attitudes, knowledge, or skills of the trainees move with the individual as they are put in different situations and teams, suggesting that CRM operates at the level of individuals. Familiarity built among team members, however, would disappear if a team does not stay together, so some benefits of CRM may be considered to reside at the team level.

We first listed each of the identified effects in each paper, and categorised these according to 'individual', 'team', or 'other'. Effects that could be placed in the individual or team category were categorised into the 'other' level, as were changes to procedures, the work environment, or 'culture'.

Across the CRM literature, multiple stances regarding compliance can be found. Helmreich et al. [17] see both intentional and unintentional incompliance with procedures as potential causes for accidents, and something that CRM training can help avoid and manage. On the other hand, Flin et al. [18] say that CRM is about non-technical skills, which are not directly related to standard operating procedures. Haller and Stoelwinder [19] go even further and argue that CRM is an alternative to a focus on operating procedure compliance. In most cases papers do not explicitly state their view on compliance. Because of that, we reduced the analysis to three categories: whether papers mention compliance should increase (or violations decrease), whether papers mention that CRM is not related to standard operating procedures, or whether there is no explicit mention of compliance to standard operating procedures.

Scope of Crew Resource Management

Finally, we considered the scope of CRM training. CRM training in the cockpit does not always focus on the same elements. For example, Salas et al. [20] mention that CRM is about coordination, and say CRM is about improving team competencies. Helmreich, Klinect, and Wilhelm [17], on the other hand, say that improving teamwork is important, but not the primary goal of CRM. To assess the scope of CRM programs in different industries, we analysed papers based on their claims of what type of errors CRM helps manage, and on the modules included in the program.

For error specification, we analysed all papers for mentions of reducing error. We included all forms of reduction, including preventing, trapping, mitigating, or handling of errors. Besides error, we included failures and mistakes. For specification, we divided this into human error, team error, and technical fault where this was specified. For team error, we included teamwork and communication errors, as these can only be observed in relation between multiple team members.

To analyse module names, we considered only papers that reported the modules names in a CRM program, and combined different papers that reported the same CRM program. As there is no standardisation in modules names, we categorised the modules into the NON-TECHNICAL Skills behavioural marker system (NOTECHS) [21] skills to allow comparison. We used identifiers based on description of each skill in the NOTECHS report to recognise similar skills (Table 1). We added 'personal resources' and an 'other' category to the NOTECHS list. 'Personal resources' was included because this was mentioned in the original NOTECHS report, but excluded from the final document because it was considered too hard to observe for a behavioural marker list. The 'other' category was added because some training modules might not fit within the NOTECHS classification. If a module fitted within two NOTECHS skills, it was counted as half a point for both skills.

Table 1. NOTECHS classification for crew resource management (CRM) modules.

NOTECHS	Identifiers
Decision Making	Problem solving
Situation Awareness	Situational awareness, situation assessment
Co-operation	Communication, teamwork, collaboration, feedback, influencing, conflict resolution, exchange of objections, challenge and response
Leadership and Managerial	Authority, assertiveness, leadership/following, obedience, maintaining standards, planning, briefings, coordination
Personal resources	Stress (coping, management), fatigue, crisis psychiatry
Other	Any non-introduction module that did not fit into the other categories

Synthesis

For the synthesis of conceptualisation data items, we combined the single paper analyses into counts per industry and percentage of industry total. For CRM modules, the total per industry only reflects papers that specified training modules. In addition, since multiple modules can cover the same skill, we calculated how many CRM training programs covered a skill at all, and how much

of the training was dedicated to each skill per industry on average. To calculate how much training was dedicated to a skill, we calculated what percentage of modules within each training program fell within each NOTECHS skill, and averaged this out for all studies that specified modules per industry.

Bias

For selecting papers, the primary potential for bias with reference to the research questions was including or excluding studies based on the authors' conceptualisation of CRM. To mitigate this, the selection process did not make judgements about what was or was not CRM—if a paper described a program labelled using one of the industry search terms, it was considered to be CRM.

On the level of individual papers, there can be a bias in academic papers for the training programs to sound more similar than the training programs actually were. Academic publications favor literature reviews with references to other academic articles. This makes it likely for researchers use words and descriptions used in other articles, even if they do not perfectly describe the studied training program. To minimise this, we have included non-academic papers and excluded papers that did not make first-hand comments.

We are unable to analyse the potential for publication bias, but it is reasonable to assume that the academic literature overrepresents novelty in CRM programs, and underrepresents programs that are very similar to each other.

2.4. Evaluation Review

2.4.1. Analysis

All papers were analysed based on their formal outcome measures, as well as other comments they made about the effects of CRM programs. This allowed us to qualitatively assess the nature of the claimed success of CRM, and follow this up by an assessment of the evidence for these claims. Something was interpreted as a claim for the success of CRM if it asserted a link between an outcome effect and the CRM program described in the study. For something to be counted as evidence for the effectiveness of CRM, we used the following evaluation criteria:

- Makes a comparison, either before/after the training or with control group;
- Tests to rule out random variation and reports basic statistical information, including means, effect sizes, and P values;
- Reports on how the measurements were made, and in the case of questionnaires, uses validated questionnaires. While ideally questionnaires are validated per type of operator and per industry, we accepted questionnaires if they were based on a validated questionnaire;
- Analyses the data according to the theoretical constructs they aim to measure.

2.4.2. Synthesis

For the synthesis of the evaluation data, we ordered the claims and evidence into a grid sorted by industry, evaluated variable, and direction of the results. We categorised the effects of CRM training using the Kirkpatrick levels of training evaluation [22]. However, as we put comparison as one of the requirements for evidence, we left out the reaction level, as reactions are not evaluated in comparison to anything. Theoretically, it would be possible to go beyond the Kirkpatrick levels, and look at industry wide changes in accident causation after the introduction of CRM. However, the subjectivity involved in accident cause determination [23–25] makes it impossible to meaningfully review this level and is beyond the scope a review on how CRM has been adopted.

2.4.3. Bias

Our evaluation could be biased towards showing CRM as more often effective as it usually is. Likely only more thoroughly developed CRM programs are going to be evaluated, which filters out

many less thorough and less well-funded CRM programs. In addition, evaluations that show CRM as effective are more likely to be published than studies that did not find CRM to be effective. This means we expect to review the better CRM programs of the reviewed industries. This does not affect our research question directly, as we are interested in how CRM has been adopted rather than its efficacy per se, but it does mean that our summary of how CRM has been evaluated may misrepresent how effective CRM has been on average.

3. Results

Figure 1 shows the numbers of search results and studies screened, assessed, and included for different parts of the review, for each industry. During the screening process, most studies were excluded because the records were not about the relevant industry. One report from MAR was excluded because the full text could not be located, but based on the description, another publication of the same study has been included. During the full-text assessment, most papers were excluded because they neither described a CRM program nor made suggestions on what a CRM program should look like. In MAR, three studies were excluded because of their poor quality of English. For the synthesis of the conceptualisation, 42 articles were used (MAR ($n = 18$) [26–43], NPI ($n = 8$) [44–51], O&G ($n = 10$) [52–61], ATC ($n = 6$) [62–67]). Appendix A shows the full study characteristics per study for the conceptualisation analysis. Appendix B shows the analysis reports for modules reported in CRM programs.

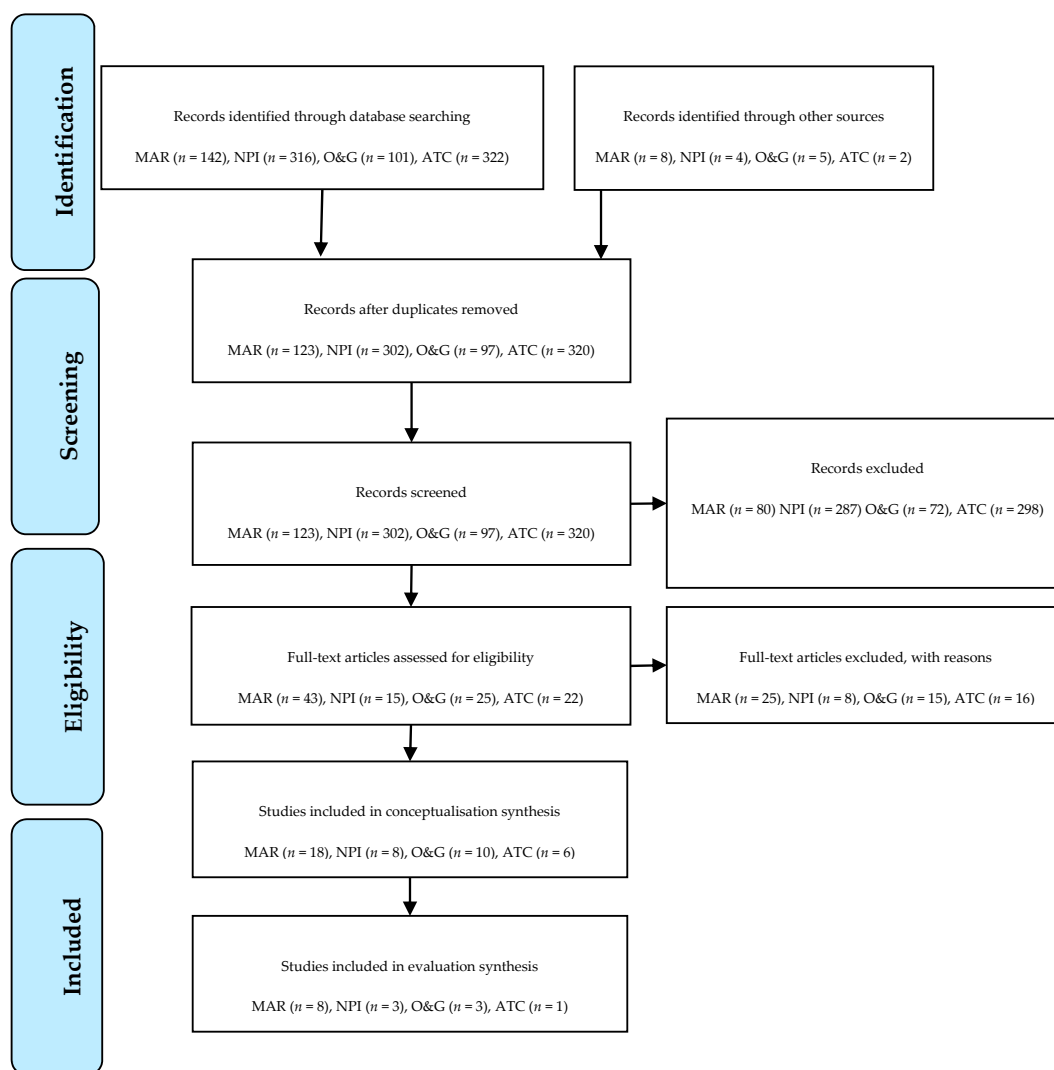


Figure 1. PRISMA flow diagram.

3.1. Analysis and Synthesis for Conceptualisation of Crew Resource Management

CRM programs are conceptualised in terms of goals (linked to the larger operational system in an industry), the levers of change, and the scope of the programs. Between programs and industries there is agreement on the broad goals, but a wide variety of levers of change. The scope of programs is varied between industries, with some uniformity within each industry.

3.1.1. Organisational Goals

Lauber [15] states that the goals of CRM are to achieve safe and efficient operations. We analysed whether the goals of ‘safety’ and ‘efficiency’ were returned in all analysed industries. Table 2 shows the percentage of papers per industry that mention either of these goals in relation to CRM training. We interpreted safety as an avoidance of negative events (accidents and incidents) and translated efficiency into more general performance improvement, as efficiency might be hard to apply to certain industries.

Table 2. Percentage (number) of papers that mention the specified organisational goal per industry.

Goal	MAR (18)	NPI (8)	O&G (10)	ATC (6)	All (42)
Safety	94% (17)	100% (8)	100% (10)	100% (6)	98% (41)
Performance	78% (14)	100% (8)	60% (6)	100% (6)	81% (34)

The papers link CRM training to the safety goals and to the performance goals of an organisation in all industries. All but one paper mention improving safety as a goal (98%) (The exception is a paper from the maritime industry investigating the relation between shared mental models, CRM, and performance [32], arguably making safety not the topic of the paper, but still a goal of the training program). Most of the papers also make the link between CRM and improving performance (81%). There are, however, some differences between industries in this regard: in the NPI and ATC industries all papers mention performance goals, in contrast to the MAR (78%) and O&G (60%) sectors. Although this difference in percentages may reflect a difference in the emphasis of CRM objectives between industries, it can also be caused by the sample, as the papers not mentioning performance tend to be short papers.

3.1.2. Levers of Change

Not all CRM programs are designed on the same philosophy and they do not expect to reach their goals in the same way. For example, Salas et al. [20] mention that some programs focussed on attitudes and others focussed on knowledge or skills. Salas et al. also commented that this difference is often hard to distil from papers. To assess differences in the envisioned levers by which CRM affects the organisational goals, we analysed papers for indications of their unit of analysis, and whether increased compliance was a mechanism to achieve the larger goals.

Unit of Analysis

All papers were analysed on the units of analysis behind desired or claimed effects of CRM training. Table 3 shows what percentage of papers mention effects on the individual level, the team level, and other levels of change.

Table 3. Percentage (number) of papers that link CRM to effects at the individual and team level.

Unit of analysis	MAR (18)	NPI (8)	O&G (10)	ATC (6)	All (42)
Individual level	100% (17)	100% (8)	100% (10)	100% (6)	100% (42)
Team level	17% (3)	0 % (0)	20% (2)	17% (1)	14% (6)
Others	6% (1)	38% (3)	60% (6)	33% (2)	29% (12)

The analysed papers mention team effects much less frequently than individual effects (14% versus 100%) in relation to CRM training across all industries (0–20% for team effects). Examples of team level effects of CRM include team bonding [59], team-member specific knowledge [66], and shared mental models [32].

In terms of other changes linked to CRM, we see bigger differences between industries. In MAR (6%) few ‘other’ effects are mentioned, while in O&G (60%) most of the papers make mention of effects that do not directly reside in individuals or teams. Examples of changes at other levels include: culture [47,56], procedure changes [52], (practical) changes to work environment [47,51], aligning teams with each other [59], testing out plans [58], and discovering training requirements [40]. These do not all have the same unit of effect. Changes in culture or procedures are expected to affect new workers, even if they have not had CRM training. The benefits from aligning teams only remain if the same teams continue interacting with each other.

Overall, there are different ideas about how CRM achieves the larger goals. All papers agree that there are changes on the individual level, but a portion of the papers suggest changes that reside at other levels of analysis. These, however, do not follow clear tendencies between industries, as many of the mentioned changes are unique. CRM programs are diverse.

Compliance

The papers were analysed on the question of whether CRM achieves its goals because it leads to better procedural compliance. Table 4 documents the percentages of papers per industry that link CRM to compliance with procedures, or that state CRM is not about procedures.

Table 4. Percentage (number) of papers that link CRM to compliance to procedures or that state CRM is beyond procedures.

Procedure relation	MAR (18)	NPI (8)	O&G (10)	ATC (6)	All (42)
Compliance	33% (6)	25% (2)	10% (1)	17% (1)	24% (10)
Not about procedures	28% (5)	38% (3)	10% (1)	0% (0)	21% (9)
Not specified	39% (7)	38% (3)	80% (8)	83% (5)	55% (23)

Less than a quarter of the papers (21%) explicitly mention that CRM training is not about operating procedures. There is some variation between industries (0–38%), with ATC scoring lowest. There is an approximately equal amount of papers that do link CRM to adherence to procedures (24%). Within an industry, the difference is never greater than one paper. However, most papers do not discuss whether there is a relationship between CRM and procedures (55%), especially in O&G and ATC (80–83%) sectors. Two papers, one from O&G [52] and one from MAR [40], mention that CRM training is used to change procedures, but these seem to be exceptions. Overall, there are differences between CRM programs, but there are no tendencies between industries.

3.1.3. Scope of Crew Resource Management

Because CRM does not always focus on the same elements, we compared different industries based on their different program compositions. To assess the scope of CRM programs in different industries, we analysed papers on what type of errors CRM helps manage and the modules included in the program.

Error Specification

For flight crews, CRM was introduced in response to errors that were observed in accidents. However, this approach does not have to be universal, nor do programs necessarily focus on all error types. Table 5 shows the percentages of papers per industry that mention the reduction of error as a goal in relation to CRM training. Furthermore, the table shows the specification of the types of errors CRM addresses, as an indication of where the emphasis lies in different sectors.

Table 5. Percentage (number) of papers that mention human, team, or technical errors per industry.

Error Type	MAR (18)	NPI (8)	O&G (10)	ATC (6)	All (42)
Human	72% (13)	50% (4) *	60% (6)	50% (3)	64% (27)
Team	16% (3)	63% (5)	50% (5)	67% (4)	40% (17)
Technical	0% (0)	25% (2)	30% (3)	0% (0)	14% (6)
Errors, all	83% (15)	75% (6)	80% (8)	100% (6)	83% (34)

* In NPI, one more paper mentions human error in its introduction, but concludes that CRM only affects team errors, not human errors. This paper therefore is only counted as team error and not as human error.

The papers generally link CRM training to a reduction in errors (83%). In ATC (100%) all papers mention errors in relation to CRM; in the other industries this percentage is slightly lower (75–83%). Also, as expected, technical errors are mentioned in only a few papers on CRM: between 0 and 30%, with an average of 14%. In contrast, a reduction in both human and team errors are associated with CRM training in half or more of the papers in each industry, except for team errors in maritime (16%). CRM is strongly related to a reduction in team errors in the NPI and ATC industries (63% and 67% respectively, in each case larger than the score on human error of 50%). Note that there is some ambiguity in classification: some authors see team errors as a subset of human errors [56], while others suggest they are separate categories [51]. These differences in categorisation make it hard to draw conclusions, but it seems that NPI and ATC put the most emphasis on the team aspects, and in O&G the team aspect is only one of the addressed aspects.

Course Modules

To look at the scope of CRM programs, we searched for the modules included in training programs. To be able to compare industries, we categorised all modules along the NOTECHS skill list [21]. Table 6 shows the percentages of programs that include a module that covers (part of) a NOTECHS skill, as well as what percentage of the modules within a program are about each skill on average per industry.

Table 6. Modules in CRM training by NOTECHS categorisation.

Module	MAR (11)		NPI (2)		O&G (5)		ATC (2)		All (20)	
	Present in % of All Prog	Avg % of Prog.	Present in % of All Prog	Avg % of Prog.	Present in % of all Prog	Avg % of Prog.	Present in % of All Prog	Avg % of Prog.	Present in % of All Prog	Avg % of Prog.
Decision making	36%	5%	100%	23%	80%	15%	100%	22%	60%	11%
Situation Awareness	36%	5%	100%	16%	40%	7%	100%	12%	50%	7%
Leadership and management	100%	35%	100%	16%	80%	17%	100%	19%	95%	27%
Cooperation	100%	23%	100%	23%	80%	24%	100%	24%	95%	24%
Personal resources	45%	6%	0%	0%	100%	28%	100%	12%	55%	12%
Other	82%	25%	100%	23%	20%	10%	50%	11%	65%	11%

The variation in CRM programs resembles the conflict in the theoretical literature. The skills ‘decision making’ (60%) and ‘situation awareness’ (50%), are reflected only in about half of all programs. All of the analysed programs in NPI and ATC had modules on these skills; however, both only have two papers that described the modules in the program.

‘Leadership and management’ (95%) and ‘cooperation’ (95%), the skills most clearly related to the team element, are covered in practically all CRM training programs. Only one CRM training program from O&G did not have modules that would fall into these two skills. When looking at how much of each training was dedicated to each skill, it becomes even clearer that these skills receive the most emphasis, as programs often had multiple modules that fall within these two skills. In most industries, more modules cover cooperation elements than leadership and management, with the

exception of MAR. This could reflect a more clearly defined hierarchy and role distribution in MAR than in other fields.

Modules related to personal resources are either about stress or fatigue. The original NOTECHS did mention these as possible skills, but did not include these skills because they were considered too difficult to observe. Modules about personal resources (55%) are, however, as common as decision making (60%) and situation awareness (50%). Only NPI (0%) has no program that mentions this skill.

More than half of all papers (65%) include modules that do not fit into the NOTECHS 'plus personal resource' categorisation. These non-fitting modules are least common in O&G (20%), and most common in MAR (82%) and NPI (100%). When looking at the modules in the 'other' category, most are unique cases. Topics that return multiple times are: emergency management-related modules in four MAR (36%) trainings; error and accident analysis in three MAR (27%) trainings and in one NPI training; and attitudes in two MAR (18%) trainings and one ATC (50%) training. The error and attitude modules are likely framing modules, rather than representing different training objectives. That emergency management often gets its own module in MAR, suggests that there people see emergency interactions as fundamentally different, which is supported by a comment that hierarchical leadership works for normal situations on ships, but a flatter hierarchy is required in emergencies [29].

These results show that there is a general agreement that the team element is part of CRM. It is likely the case that many training programs see CRM as more than just about the team element, but the size of this group cannot be drawn from our analysis. Certain programs talk about fatigue and self-management [28], while other programs talk about how stress can be recognised in other people [66], relating stress to the team element. Just like personal resources, decision making and situation awareness can be addressed for individuals or for teams. Thus, the inclusion of personal resources, decision making, or situation awareness does not necessarily indicate CRM training aims to target non-team processes.

3.1.4. Conceptualisation Synthesis Summary

There is agreement about the goals of CRM, but there is disagreement among papers in how CRM reaches those goals and what falls within the scope of CRM training. Almost all papers mention safety and performance as goals. All papers agree that part of the effects can be found on the individual level, but a decent amount of papers talk about changes at other levels as well. Papers are split on whether CRM is or is not about compliance and procedures in all industries. In terms of scope, there seems to be a difference per industry. NPI, ATC, and MAR seem to put the most emphasis on the team element, while in O&G a broader approach is used, where human elements in general are addressed. This diversity might be larger than our analysis shows, as we expected bias towards similar language, concealing differences in conceptualisation and implementation. The diversity in how researchers conceptualise CRM, both within and between industries, needs to be reflected in the evaluation of CRM in order for CRM to be considered properly evaluated.

3.2. Analysis and Synthesis for Crew Resource Management Evaluation

In the review of the evaluation of CRM programs, we searched the papers for all claims of effects arising from the training, and then analysed how these claims held up against our evaluation criteria. Because of the low number of studies passing our criteria and for the purpose of clarity, we discuss the analysis of all studies individually, before we synthesise the results. We discuss effects at the learning, application, and organisational level per industry. This is followed by a summary to provide an overview of the evidence supporting CRM training effectiveness in the different industries.

3.2.1. Maritime

Learning

In MAR, five studies examined learning effects. Fonne and Fredriksen [27] found that attitudes changed directly after the course, and moderated over a period of six months. However, they only report percentages and did not do any statistical tests to rule out random variation.

Inoue and Takahashi [33] used one-item attitude-like measures, of which they compared the averages, and found that junior trainees changed their attitude more than senior officers. They did no statistical test on the before/after change, or between groups, to rule out random variation.

Saatcioglu et al. [35] claimed that trainees changed their attitudes towards cooperative learning. However, they tested for statistical difference per item of the questionnaire, not per the attitudes that make up cooperative learning.

Håvold et al. [41] investigated how CRM training achieved its effect. Using a questionnaire, they found that people who thought the course was better set up also believed their knowledge, skills, and attitudes changed more. In addition, those who believed their knowledge, skills, and attitudes had changed more also thought their behaviour had changed more. However, the questionnaire was only assessed for reliability, but not validity, and it was a multiple-regression study, not a comparison study.

Brun et al. [32] looked at the effects of CRM training on shared mental models. They found no difference between the experimental and control group. The teams in the study had already worked together before the training, which can explain why no change was found. Cockpit CRM has usually been tested with newly formed teams. The test was too small to expect any effect of statistical significance, so the results give no indication whether CRM has an effect on shared mental models in already established teams.

None of these studies meet our evaluation criteria.

Behaviour

Four papers looked at how CRM training was applied by the trainees and led to behavioural changes. Byrdorf [28] and Håvold et al. [41] used questionnaires and interviews to ask people whether they had changed their behaviour afterwards. In both studies, people reported changed behaviour, but neither study did a comparison, either a 'before-and-after' or with a control group. Brun et al. [32] and Wu et al. [42] both compared two teams per condition and found small differences. Brun et al. [32] doubted whether the differences were caused by CRM, while Wu et al. [42] explained this as a small effect size from CRM training. Both studies were too small for a statistical test.

None of the studies meet our evaluation criteria.

Organisational Effects

In terms of organisational goals, two MAR studies reported effects. Byrdorf [28] found big changes on multiple organisational measures when comparing 1992, the year the intervention started, with 1996, four years after they started doing the training. They found large improvements on multiple measures. However, there was no control group and only two years are reported, taking away the opportunity to rule out alternative explanations. Other explanations have some merit, as there were three other safety interventions being implemented in the company [28], and there was a general trend of reduction of accidents in the industry at that time [29].

Wang et al. [40] claimed success of their CRM training because of changes they made to their planned operations. This included changes to procedures, interface design, safety margins, the flow of information, and the identification of additional training needs. These outcomes make CRM more into a test of plans, as opposed to the development of skills. These outcomes can be as useful for both safety and efficiency, but it is impossible to verify whether these suggestions improved safety or efficiency.

Neither study meets our evaluation criteria.

3.2.2. Nuclear Power Industry

Learning

Two studies in the NPI looked at learning effects of CRM training. Harrington and Kello used a questionnaire based on the Controller Resource Management Attitude Questionnaire (CMAQ) and found an effect on all of its attitudes, recognition of stressor effects, communication and coordination, and command responsibility.

Kim and Byun used the questionnaire based on the Flight Management Attitude Questionnaire (FMAQ) and found a statistically significant change in the desired direction. However, they only tested for all items grouped together and did not analyse the data per attitude, making it unclear what the significant difference represents.

The results of Harrington and Kello [45] meet our evaluation criteria and support attitude change from CRM, but the results of Kim and Byun [51] on attitudes do not meet our evaluation criteria.

Behaviour

Kim and Byun [51] compared a CRM-trained group and a control group on multiple measures in a simulator. They measured situation awareness and workload through self-report, for which they found no effects. For team behaviour, they used behavioural markers using evaluators and found a significant difference there between groups. Their conclusion was that CRM mostly works at the team level, not individual skills. However, the groups were not randomly selected; the CRM-trained group was more experienced and familiar with operating a different type of nuclear reactor. In addition, the evaluators were also the designers and trainers of the course, which could have biased the ratings. This evaluation meets our evaluation criteria, although alternative explanations for the effects do exist.

Organisational Effects

Davis [47] found that a CRM program led to improvement in operational systems, attitudes, culture, human performance, and reportable events associated with human performance and supervision. However, the numbers are not reported, nor does he explain how they were measured. With that, the study does not meet our evaluation criteria.

3.2.3. Oil and Gas

Learning

One study in O&G evaluated learning effects. O'Connor and Flin [57] looked at the learning in terms of attitudes and knowledge of the trainees before and after the training. For attitudes they used a CMAQ based questionnaire, which they reduced from the original six factors of the questionnaire to four. Of these four factors, decision making and personal limitations changed in the desired direction after the training, but there was no change for situation awareness and communication. The researchers reported multiple issues with their test, which can include explanations for the lack of effect of the training, (1) because of the high initial scores; (2) the small sample size compared to the expected effect size; and (3) the reliability of the scales was not great. In their knowledge test they found no changes. A possible reason given for this was that practitioners might not have been that motivated for the last test, as they were about to go home. The study meets our selection criteria and provides mixed support for CRM training changing attitudes.

Behaviour

The only study that evaluated CRM training on the level of application comes from Moffat and Crichton [61]. They looked at behaviour change from teams during multiple exercises in CRM training. They observed changed scores between the earlier and final exercise. However, the data is limited to two exercises, and the marker system was still in development, meaning the scale had not been

tested on reliability or validity. It is unclear whether the evaluators were also the trainers of the course. The study can be expanded, but in its current form the study does not meet our selection criteria.

Organisational Effects

For organisational effects, Crichton [58] mentioned that a trained team had an exemplary performance in terms of non-productive time, health and safety data, and project goals. However, the measures are not reported and the sample size was 1. The study sounds promising, but the study does not meet our requirements for verification.

3.2.4. Air Traffic Control

Learning

There was only one paper that evaluated CRM training in ATC. It was a larger EUROCONTROL project that tested a program that was rolled out in multiple European countries [65]. This study used an FMAQ-based questionnaire to test for attitudes before and after the training. For all items combined they found a statistically significant change. On another smaller sample they found statistically significant change on 17 of the 38 items; however, some of these items changed in the undesired direction. The study lacks a comparison per attitude, which the items are supposed to reflect. Because of that, the study does not meet our selection criteria.

3.2.5. Evaluation Synthesis Summary

A summary of the evaluation in different industries can be found in Table 7. There are six evaluations that meet our quality criteria, of which two are positive, one is mixed, and three show no effect. This means that in the most rigorous tests, CRM is effective less than half of the time. Both tests with positive results come from NPI, but NPI also has two tests showing no effect. In O&G there is one test with mixed results and one test that shows no effect. In MAR and ATC none of the analysed studies meet our criteria. In all of the cases that meet our evaluation criteria, the measurements were done right after the training. This leaves the question whether any of the few found changes are lasting.

There are many more claims about the effectiveness of CRM that do not meet our evaluation criteria, and that are generally more positive. Some of these claims are from short or explorative papers, which might have deliberately traded off between costs and thoroughness of the evaluation. However, there are also papers where this explanation does not fit, such as when papers separately analyse individual items intended to form a single dimension, or combine multiple dimensions designed to be separate. We do not know why studies would do this.

Table 7. Summary of evaluation claims and results.

Evaluation	Status	MAR			NPI			O&G			ATC		
		Pos	Mix	No	Pos	Mix	No	Pos	Mix	No	Pos	Mix	No
Attitudes	Tested	-	-	1	-	-	-	1	-	-	-	-	-
	Claim	4	-	-	3	-	-	1	-	-	1	-	-
Knowledge	Tested	-	-	-	-	-	-	-	-	1	-	-	-
	Claim	1	-	-	-	-	-	-	-	1	-	-	-
Skills	Tested	-	-	-	-	-	-	-	-	-	-	-	-
	Claims	1	-	-	-	-	-	-	-	-	-	-	-
Shared mental models	Tested	-	-	-	-	-	-	-	-	-	-	-	-
	Claims	-	-	1	-	-	-	-	-	-	-	-	-
Situation Awareness	Tested	-	-	-	-	-	1	-	-	-	-	-	-
	Claims	-	-	-	-	-	1	-	-	-	-	-	-
Workload	Tested	-	-	-	-	-	1	-	-	-	-	-	-
	Claims	-	-	-	-	-	1	-	-	-	-	-	-
Behaviour	Tested	-	-	1	-	-	-	-	-	-	-	-	-
	Claims	3	1	1	-	-	-	1	-	-	-	-	-
Organisational	Tested	-	-	-	-	-	-	-	-	-	-	-	-
	Claims	2	-	1	-	-	-	1	-	-	-	-	-

There is no indication that the different conceptualisations of CRM are taken into account in the evaluation measures. We have seen no mention of compliance. In the behavioural measures, there does not seem to be a distinction that O&G focuses more on individual aspects than other industries. There is one team level concept measured, shared mental models, which is less than the 14% of the papers that mention team level effects. Papers do mention changes that do not reside at the individual or team level, such as changes to rules, procedures, training requirements, and climate, but none of these evaluations meet our evaluation criteria. In addition, we have not found any evidence of relations between any of the evaluation elements, which could have favoured one conceptualisation over another.

4. Discussion

4.1. Summary of Evidence

CRM has spread to new domains. There is, however, no unified view of what CRM is or how it is supposed to improve safety and efficiency. There are some similarities in the scope of what CRM programs encompass within industries, but in terms of levers of change there is as much diversity within industries as between industries. Many studies claim that a CRM program has been effective, and offer limited evidence to support these claims. Many studies do not rule out random variation, or do not analyse questionnaire data consistently with the dimensional structure of the questionnaire. In addition, the conceptualisation of CRM across studies does not align with how studies evaluate CRM, leaving a large part of what researchers say about CRM unevaluated.

4.2. Limitations

The largest identified biases in this study should point towards the finding that CRM programs are more similar than they actually are, and that CRM is more effective than it actually is. In fact, the review results pointed in the opposite direction to these biases.

It is still possible that the academic literature does not capture the full diversity of CRM in the reviewed industries. There may be really successful CRM training programs in industries that are not reported academically, but it is more likely that on average unreported CRM training is less effective, as we expect that the better programs are the ones that are evaluated and published.

For the conceptualisation, we have relied on the description authors made at the time of publication. We did not seek out current training material from the programs. There may be changes over time regarding the conceptualisation and scope of CRM programs. Our approach gave us a wider view and allowed us to incorporate more evidence, but this means that the results do not necessarily reflect the most current views in an industry.

4.3. Conclusions

Previous reviews of CRM have argued for more evaluation, especially across multiple levels (individual, team, and organisational effects). We agree that showing a decrease in accidents is one useful outcome measure, but this is probably not enough. Evaluation needs to focus on the process through which CRM programs lead to change. This is not something that can be assumed from the label 'CRM', because of the differences between programs in how CRM is conceptualised. Each CRM study needs to make their hypotheses about how CRM creates change explicit. Salas et al. [4] argued for standardisation in the training and evaluation of CRM. That might have made sense once, but considering CRM's long history and diverse fields of application, we now argue for the opposite. We suggest researchers go behind the label CRM and describe what they intend to do and expect to change. This can make the differences that are normally hidden behind the label CRM visible and create a more auspicious way to evaluate the effectiveness of CRM in all its diversity.

For example, there is good indication that the training model that is reflected in the Kirkpatrick evaluation levels does not work well for CRM. This model suggests that the training leads to learning with changes in knowledge, skills, and attitudes, which leads to a change in behaviour, leading

to improved safety and performance. Of all these elements, it is most supported by research that CRM leads to changes in attitudes and behaviour, but the links between these elements are not well established. Two studies have touched on the link between behaviour and attitudes, but do not provide strong evidence for this link. The oldest paper was about pilots [68] and has methodological shortcomings. First, it looks at separate items of a questionnaire, not the attitudes they represent. Second, the predictive values are based on a test with the same dataset from which the model was built, instead of testing it on a new dataset. A more recent paper from the maritime sector [69] found no link between attitudes and behaviour, but further exploration of the data showed there might be a relation between low-mid attitudes and behaviour. As far as there was a link between attitudes and behaviour in this study, it was a universal link of all attitudes to all behaviour, not a link of specific attitudes to related behaviours. This can suggest that any link may be achieved through a single overarching variable, such as psychological safety climate [70], and not through a change in attitudes.

This leaves two possibilities. The first is that there is some link between behaviour and attitudes, but it is a weak one—considering that in all reviewed studies, the largest effect on a single attitude found was 9%, less than 0.5 points on a five-point Likert scale. When that is combined with a weak link to behaviour, it is unlikely that CRM meaningfully changes behaviour. It is possible that the claims about the success of CRM do reflect real change, but that it is not captured by this model and that it comes about through other processes, such as change in the psychological safety climate. In either case, this means that CRM training programs focus on elements that have weak effects, and that training could be optimised by spending time on different elements. To find answers to these questions, however, CRM researchers need to explicitly describe how they expect change to take place after the training and test accordingly.

Acknowledgments: This research is part of a larger research project funded by ORICA. The views presented are those of the authors and should not be taken to represent the position or policy of the funding bodies.

Author Contributions: All authors were involved in designing the review process and defining the categories. Jop Havinga conducted the coding of the papers and analysing the results. All authors were involved in the writing of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Characteristics of studies meeting selection criteria for conceptualisation review.

Author	Year	MAR			
		Goals	Level of Effect	Compliance	Errors
Fonne and Fredriksen [27]	1995	S, P	I	C	T
Dijkhuizen et al. [26]	1996	S	I	n/s	H
Byrdrof [28]	1998	S, P	I	C	H
Habberley et al. [29]	2001	S, P	I	B	H, T
Barnett et al. [30]	2003	S, P	I, T	B	H
Kim [31]	2002	S, P	I	C	H
Brun et al. [32]	2005	P	I, T	B	
Inoue and Takahashi [33]	2005	S, P	I	n/s	H
Pekcan et al. [34]	2005	S, P	I	B	H, T
Saatcioglu et al. [35]	2005	S, P	I	n/s	E
Barsan et al. [36]	2010	S	I	n/s	
Lee et al. [37]	2010	S	I	C	H
Thi et al. [38]	2012	S, P	I	n/s	H
Dunham and Lützhöft [39]	2014	S, P	I	n/s	
Wang et al. [40]	2014	S, P	I, T, O	C	H
Håvold et al. [41]	2015	S, P	I	B	H
Wu et al. [42]	2015	S, P	I	n/s	H
Deboo [43]	n.d.	S, P	I	C	H, T

Table A1. Cont.

NPI					
Author	Year	Goals	Level of Effect	Compliance	Errors
Harrington and Kello [44]	1991	S, P	I	n/s	
Harrington and Kello [45]	1992	S, P	I	n/s	
Gaddy and Wachtel [46]	1992	S, P	I	n/s	T
Davis [47]	2001	S, P	I, O	C	H
Crichton and Flin [48]	2004	S, P	I	B	H, T
O'Connor et al. [49]	2008	S, P	I	C	H, T
Kim et al. [50]	2009	S, P	I, O	B	H, T, Tn
Kim and Byun [51]	2011	S, P	I, O	B	T, Tn
O&G					
Author	Year	Goals	Level of Effect	Compliance	Errors
Grinde [52]	1994	S	I, T, O	n/s	
Flin [53]	1995	S, P	I	n/s	E
Flin [54]	1999	S, P	I	n/s	H, T
Flin and O'Connor [55]	2001	S, P	I	n/s	H, T, Tn
Mearns et al. [56]	2003	S, P	I, O	n/s	H, T, Tn
O'Connor and Flin [57]	2003	S, P	I, O	n/s	H, T, Tn
Crichton [58]	2009	S, P	I, O	n/s	E
Bailey [59]	2014	S	I, T, O	n/s	
Thorogood and Crichton [60]	2014	S	I, O	C	H
Moffat and Crichton [61]	2015	S	I	B	H, T
ATC					
Author	Year	Goals	Level of Effect	Compliance	Errors
Henderson [62]	1988	S, P	I	C	H, T
Woldring [63]	1999	S, P	I	n/s	T
Barbarino and Isaac [64]	1999	S, P	I, O	n/s	T
Woldring and Isaac [65]	2000	S, P	I, O	n/s	T
Smith-Jentsch et al. [66]	2001	S, P	I, T	n/s	H
Park et al. [67]	2006	S, P	I	n/s	H

Key: Goals: S = Safety; P = Performance. Level of effect: I = Individual; T = Team; O = Other. Compliance: C = Compliance; B = Beyond procedures; n/s = not specified. Errors: E = Errors unspecified; H = Human error; T = Team error; Tn = Technical error.

Appendix B

Table A2. CRM Program module analysis.

Author	Year	Industry	Modules					
			DM	SA	LM	CP	PR	Other
Dijkhuizen et al. [26]	1996	MAR	0	0	3.5	0.5	0	2
Byrdrof [28]	1998	MAR	0	0	3	2	1	0
Kim [31]	2002	MAR	0	1	3.5	3	0.5	0
Brun et al. [32]	2005	MAR	2	1	1	1	1	1
Pekcan et al. [34]	2005	MAR	0	1	1	2	0	3
Saatcioglu et al. [35]	2005	MAR	0	1	3	2	0	3
Lee et al. [37]	2010	MAR	1	0	2	2	0	4
Thi et al. [38]	2012	MAR	0.5	0	1.5	2	1	0
Håvold et al. [41]	2015	MAR	0	0	2	2	0	4
Wu et al. [42]	2015	MAR	0	0	3	1	1	2
Deboo [43]	n.d.	MAR	1	0	7	2	0	4
Davis [47]	2001	NPI	2	1	1	2	0	2
Kim and Byun [50,51]	2011	NPI	1	1	1	1	0	1
Grinde [52]	1994	O&G	0	0	0	0	1	1
Flin [53,54,56]	1995	O&G	1	0	1	1	1	0
Flin [53,54,56]	1995	O&G	1	0	2	3	1	0
Mearns et al. [56,57]	2003	O&G	1	1	1	1	2	0
Moffat and Crichton [61]	2015	O&G	1	1	1	2	1	0
Henderson [62]	1988	ATC	4	1	3	2	1	4
Woldring and Isaac [65]	2000	ATC	1	1	1	2	1	1

Key: DM = decision Making; SA = Situation Awareness; LM = Leadership and management; CP = cooperation; PR = Personal Resources.

References

- Maurino, D.E.; Murray, P.S. Crew resource management. In *Handbook of Aviation Human Factors*, 2nd ed.; Wise, J.A., Hopkin, D., Garland, D.J., Kantowitz, B., Eds.; CRC Press: Boca Raton, FL, USA, 2009.
- Kanki, B.; Helmreich, R.; Anca, J. (Eds.) *Crew Resource Management*; Elsevier: Amsterdam, The Netherlands, 2010.
- Salas, E.; Wilson, K.A.; Burke, C.S.; Wightman, D.C.; Howse, W.R. Crew Resource Management Training Research, Practice, and Lessons Learned. *Rev. Hum. Factors Ergon.* **2006**, *2*, 35–73. [[CrossRef](#)]
- Salas, E.; Wilson, K.A.; Burke, C.S.; Wightman, D.C. Does crew resource management training work? An update, an extension, and some critical needs. *Hum. Factors* **2006**, *48*, 392–412. [[CrossRef](#)] [[PubMed](#)]
- Edkins, G.D. A review of the benefits of aviation human factors training. *Hum. Factors Aerosp. Saf.* **2002**, *31*, 247–273.
- Holt, R.W.; Boehm-Davis, D.A.; Beaubien, J.M. Evaluating resource management training. In *Improving Teamwork in Organizations: Applications of Resource Management Training*; Salas, E., Bowers, C.A., Eden, E., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 2001; pp. 165–188.
- Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. The PRISMA GROUP Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, 1–6. [[CrossRef](#)] [[PubMed](#)]
- O'Connor, P.; Flin, R.H.; Flechter, G. Methods used to evaluate the effectiveness of CRM training: A literature review. *J. Hum. Factors Aerosp. Saf.* **2002**, *2*, 217–234.
- O'Connor, P.; Campbell, J.; Newon, J.; Melton, J.; Salas, E.; Wilson, K.A. Crew Resource Management Training Effectiveness: A Meta-Analysis and Some Critical Needs. *Int. J. Aviat. Psychol.* **2008**, *18*, 353–368. [[CrossRef](#)]
- O'Dea, A.; O'Connor, P.; Keogh, I. A meta-analysis of the effectiveness of crew resource management training in acute care domains. *Postgrad. Med. J.* **2014**, *90*, 699–708. [[CrossRef](#)] [[PubMed](#)]
- Buljac-Samaradzic, M.; Dekker-van Doorn, C.M.; van Wijngaarden, J.D.H.; van Wijk, K.P. Interventions to improve team effectiveness: A systematic review. *Health Policy* **2010**, *94*, 183–195. [[CrossRef](#)] [[PubMed](#)]
- Weaver, S.J.; Lyons, R.; DiazGranados, D.; Rosen, M.A.; Salas, E.; Oglesby, J.; Augenstein, J.S.; Birnbach, D.J.; Robinson, D.; King, H.B. The anatomy of health care team training and the state of practice: A critical review. *Acad. Med.* **2010**, *85*, 1746–1760. [[CrossRef](#)] [[PubMed](#)]
- Salas, E.; Bowers, C.A.; Cannon-Bowers, J.A. Military Team Research: 10 Years of Progress. *Mil. Psychol.* **1995**, *7*, 55–75. [[CrossRef](#)]
- O'Connor, P.; Hahn, R.; Salas, E. The U.S. Navy's Crew Resource Management Program: The Past Present, and Recommendations For the Future. In *Human Performance Enhancements in High-Risk Environments: Insights Developments, and Future Directions from Military Research*; ABC-Clio: Santa Barbara, CA, USA, 2010; pp. 91–105.
- Lauber, J. Resource management in the cockpit. *Air Line Pilot* **1984**, *53*, 20–35.
- Flin, R.H.; O'Connor, P.; Crichton, M.T. *Safety at the Sharp End*; Ashgate Publishing Limited: Farnham, UK, 2008.
- Helmreich, R.L.; Klinec, J.R.; Wilhelm, J.A. Models of Threat, Error, and CRM in Flight Operations. In Proceedings of the Tenth International Symposium on Aviation Psychology, Columbus, OH, USA, 3–6 May 1999; pp. 677–682.
- Flin, R.H.; Martin, L.; Goeters, K.; Hörmann, H.; Amalberti, R.; Nijhuis, H. Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Hum. Factors Aerosp. Saf.* **2003**, *3*, 95–117.
- Haller, G.; Stoelwinder, J. Is "Crew Resource Management" an Alternative to Procedure-Based Strategies to Improve Patient Safety In Hospitals? In *Trapping Safety into Rules: How Diserable or Avoidable Is Proceduralization?* Bieder, C., Bourier, M., Eds.; Ashgate Publishing: Farnham, UK, 2013.
- Salas, E.; Prince, C.; Bowers, C.A.; Stout, R.J.; Oser, R.L.; Cannon-bowers, J.A. A Methodology for Enhancing Crew Resource Management Training. *Hum. Factors* **1999**, *41*, 161–172. [[CrossRef](#)]
- Van Avermaete, J.A.G.; Kruijssen, E.A. NOTECHS: *The Evaluation of Non-Technical Skills of Multi-Pilot Aircrew in Relation to the JAR-FCL Requirements*; National Aerospace Laboratory: Amsterdam, The Netherlands, 1998.
- Kirkpatrick, J.D. *Kirkpatrick's Four Levels of Training Evaluation*; ATD Press: Alexandria, VA, USA, 2016.
- Rasmussen, J. Human error and the problem of causality in analysis of accidents. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1990**, *327*, 449–462. [[CrossRef](#)] [[PubMed](#)]
- Lundberg, J.; Rollenhagen, C.; Hollnagel, E. What-You-Look-For-Is-What-You-Find—The consequences of underlying accident models in eight accident investigation manuals. *Saf. Sci.* **2009**, *47*, 1297–1311. [[CrossRef](#)]

25. Wrigstad, J.; Bergström, J.; Gustafson, P. One event, three investigations: The reproduction of a safety norm. *Saf. Sci.* **2017**, *96*, 75–83. [[CrossRef](#)]
26. Dijkhuizen, H.A.; den Butter, P.; Koning, B.R. Instruction, training and expertise advancement of pilots: The key elements in obtaining the highest possible quality of certified pilots in the Netherlands. In *Marine Simulation and Ship Manoeuvrability*; Chislett, M.S., Ed.; A.A. Balkema: Rotterdam, The Netherlands, 1996; pp. 125–143.
27. Fonne, V.; Fredriksen, O. Resource management and crew training for HSV-navigators. In Proceedings of the 8th International Symposium of Aviation Psychology, Columbus, OH, USA, 24–27 April 1995; Jensen, R.S., Rakovan, L.A., Eds.; pp. 585–590.
28. Byrdorf, P. Human factors and crew resource management: An example of successfully applying the experience from CRM programmes in the aviation world to the Maritime World. In Proceedings of the 23rd Conference of the European Association for Aviation Psychology, Vienna, Austria, 14–18 September 1998.
29. Habberley, J.S.; Barnett, M.L.; Gatfield, D.; Musselwhite, C.; McNeil, G. *Simulator Training for Handling Escalating Emergencies*; Warsash Maritime Centre: Southampton, UK, 2001.
30. Barnett, M.L.; Gatfield, D.; Pekcan, C. A Research Agenda in Maritime Crew Resource Management. In Proceedings of the International Conference on Team Resource Management in the 21st Century, Daytona Beach, FL, USA, 23–24 October 2003.
31. Kim, C.-J. Maritime Education and Training (MET) by ship handling simulator. *J. Korea Soc. Simul.* **2002**, *11*, 81–90.
32. Brun, W.; Eid, J.; Johnsen, B.H.; Laboerg, J.C.; Ekornas, B.; Kobbeltvedt, T. Bridge Resource Management Training Enhancing Shared Mental Models and Task Performance. In *How Professionals Make Decisions*; Montgomery, H., Lipshitz, R., Brehmer, B., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2005; pp. 183–193.
33. Inoue, K.; Takahashi, H. Quantitative Evaluation of Consciousness Improvement in BRM Training. In Proceedings of the 6th Annual General Assembly and Conference of the International-Association-of-Maritime-Universities, Malmö, Sweden, 24–26 October 2005; WIT: Southampton, UK, 2005; pp. 387–395.
34. Pekcan, C.; Gatfield, D.; Barnett, M.L. Content and context: Understanding the complexities of human behaviour in ship operation. In Proceedings of the RINA, Royal Institution of Naval Architects International Conference—Human Factors in Ship Design, Safety and Operation, London, UK, 23–24 February 2005; Royal Institution of Naval Architects: London, UK, 2005; pp. 99–107.
35. Saatcioglu, O.; Asyali, E.; Cerit, A. Cooperative Learning and Teamwork Effectiveness: Impacts of Education Period on Cadets. In Proceedings of the International Association of Maritime Universities (IAMU) 6th Annual General Assembly and Conference, Malmö, Sweden, 24–26 October 2005; Nielson, D., Ed.; WIT Press: South Hampton, UK, 2005; pp. 377–386.
36. Barsan, E.; Hanzu-Pazara, R. Human response in navigation emergency situation. *Int. J. Veh. Saf.* **2010**, *5*, 75–85. [[CrossRef](#)]
37. Lee, E.-B.; Yun, J.-H.; Jeong, T.-G. A Cyber-Training and Education Model for Tug-barge Operators. *J. Navig. Port Res. Int. Ed.* **2010**, *34*, 287–292. [[CrossRef](#)]
38. Thi, K.L.T.; Jeong, J.-S.; Jeong, J.-Y. A Study on the Effective Teaching Method for Bridge Resource Management in Korea. *J. Korean Soc. Mar. Environ. Saf.* **2012**, *18*, 563–568.
39. Dunham, R.C.; Lützhöft, M. Interim Results of a longitudinal study into the perceptions of Bridge Team Management (BTM) in pilotage waters as experienced by Deck Officers studying at the Australian Maritime College. In Proceedings of the 15th Annual General Assembly International Association of Maritime Universities, Launceston, Australia, 27–29 October 2014; pp. 118–124.
40. Wang, A.M.; Yuan, R.; Li, H.; Xie, W.; Bakker, J.S.; Pinkster, J. Virtual simulation program and its application for challenging floatover installation of Liwan 3-1 mega topsides in South China sea. In Proceedings of the International Offshore and Polar Engineering Conference, Busan, Korea, 15–20 June 2014; Volume 3, pp. 700–711.
41. Håvold, J.I.; Nistad, S.; Skiri, A.; Ødegård, A. The human factor and simulator training for offshore anchor handling operators. *Saf. Sci.* **2015**, *75*, 136–145. [[CrossRef](#)]
42. Wu, Y.; Miwa, T.; Shimamoto, K.; Uchida, M. Development of quantitative team performance evaluation method for ERM. *WMU J. Marit. Aff.* **2015**, *14*, 333–347. [[CrossRef](#)]

43. Deboo, C.K.N. *Maritime Resource Management—Based on the World Renowned SAS-BRM Course*; Academic Press: Cambridge, MA, USA, 2010.
44. Harrington, D.K.; Kello, J.E. Systematic evaluation of nuclear operator team skills training. *Trans. Am. Nucl. Soc.* **1991**, *64*, 182–183.
45. Harrington, D.K.; Kello, J.E. Systematic evaluation of nuclear operator team skills training: A progress report. In Proceedings of the Conference Record for 1992 IEEE Fifth Conference on Human Factors and Power Plants, Monterey, California, USA, 7–11 June 1992; IEEE: Monterey, CA, USA, 1992; pp. 370–373.
46. Gaddy, C.D.; Wachtel, J. Team skill training in nuclear power plant operations. In *Teams: Their Training and Performance*; Swezey, R.W., Salas, E., Eds.; Ablex: Norwood, NJ, USA, 1992; pp. 379–396.
47. Davis, J. Enhancing human performance in the control room: Sharing the experiences of British Energy. In Proceedings of the People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, Manchester, UK, 19–21 June 2001; p. 345.
48. Crichton, M.T.; Flin, R.H. Identifying and training non-technical skills of nuclear emergency response teams. *Ann. Nucl. Energy* **2004**, *31*, 1317–1330. [[CrossRef](#)]
49. O'Connor, P.; O'Dea, A.; Flin, R.H.; Belton, S. Identifying the team skills required by nuclear power plant operations personnel. *Int. J. Ind. Ergon.* **2008**, *38*, 1028–1037. [[CrossRef](#)]
50. Kim, S.K.; Park, J.Y.; Byun, S.N. Crew Resource Management Training for Improving Team Performance of Operators in Korean Advanced Nuclear Power Plant. In Proceedings of the 2009 IEEE Industrial Engineering and Engineering Management (IEEM), Hong Kong, China, 8–11 December 2009; pp. 2055–2059.
51. Kim, S.K.; Byun, S.N. Effects of Crew Resource Management Training on the Team Performance of Operators in an Advanced Nuclear Power Plant. *J. Nucl. Sci. Technol.* **2011**, *48*, 1256–1264. [[CrossRef](#)]
52. Grinde, T.A. Emergency Resource Management training. In Proceedings of the SPE Health, Safety and Environment in Oil and Gas Exploration and Production Conference, Jakarta, Indonesia, 25–27 January 1994; Society of Petroleum Engineers: Jakarta, Indonesia, 1994; pp. 413–417.
53. Flin, R.H. Crew resource management for teams in the offshore oil industry. *J. Eur. Ind. Train.* **1995**, *19*, 23–27. [[CrossRef](#)]
54. Flin, R.H.; O'Connor, P.; Mearns, K.; Gordon, R.P.E. Crew Resource Management for Offshore Production and Maintenance. In Proceedings of the Offshore Europe Oil and Gas Exhibition and Conference, Aberdeen, UK, 7–10 September 1999; Society of Petroleum Engineers: Aberdeen, Scotland, 1999.
55. Flin, R.H.; O'Connor, P. Applying Crew Resource Management on Offshore Oil Platforms. In *Improving Teamwork in Organizations: Applications of Resource Management Training*; Salas, E., Bowers, C.A., Edens, E., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2001; pp. 217–234.
56. Mearns, K.; Whitaker, S.; Flin, R.H.; Gordon, R.P.E.; O'Connor, P. *Factoring the Human into Safety: Translating Research into Practice: Crew Resource Management Training for Offshore Operations*; Health and Safety Executive: Aberdeen, Scotland, 2003; Volume 3.
57. O'Connor, P.; Flin, R.H. Crew resource management training for offshore oil production teams. *Saf. Sci.* **2003**, *41*, 591–609. [[CrossRef](#)]
58. Crichton, M.T. Improving team effectiveness using tactical decision games. *Saf. Sci.* **2009**, *47*, 330–336. [[CrossRef](#)]
59. Bailey, T. BP uses high-fidelity simulators to develop non-technical skills. *Offshore Eng.* **2014**, *39*, 32–34.
60. Thorogood, J.L.; Crichton, M.T. Threat-and-Error Management: The Connection Between Process Safety and Practical Action at the Worksite. In Proceedings of the IADC/SPE Drilling Conference and Exhibition, Fort Worth, TX, USA, 4–6 March 2014; pp. 465–472.
61. Moffat, S.; Crichton, M.T. Investigating Non-technical Skills through Team Behavioral Markers in Oil and Gas Simulation-based Exercises. *Procedia Manuf.* **2015**, *3*, 1241–1247. [[CrossRef](#)]
62. Henderson, E.D. Air traffic controller awareness and resource training. In Proceedings of the 7th Aerospace Behavioral Technology Conference: Operational Information Transfer Technology: Are We Designing for the Human Operator, Anaheim, CA, USA, 3–6 October 1988; SAE: Warrendal, PA, USA, 1988; pp. 141–144.
63. Woldring, M. Team Resource Management in European Air Traffic Control. *Air Space Eur.* **1999**, *1*, 81–84. [[CrossRef](#)]
64. Barbarino, M.; Isaac, A. Development of Team Resource Management in European air traffic control. In Proceedings of the Fourth Australian Aviation Psychology Symposium, Sydney, Australia, 16–20 March 1998; Lowe, A.R., Hayward, B.J., Eds.; Ashgate: Aldershot, UK, 2000; pp. 269–278.

65. Woldring, M.; Isaac, A. *Team Resource Management Test and Evaluation*; EUROCONTROL: Brussels, Belgium, 1999.
66. Smith-Jentsch, K.A.; Baker, D.P.; Salas, E.; Cannon-Bowers, J.A. Uncovering Differences in Team Competency Requirements: The Case of Air Traffic Control Teams. In *Improving Teamwork in Organizations: Applications of Resource Management Training*; Salas, E., Bowers, C.A., Edens, E., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2001; pp. 31–54.
67. Park, S.A.; Sohn, Y.W.; Kim, K. Shared Perception of Risks in Emergency Situations between Pilots and Air Traffic Controllers. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Francisco, CA, USA, 16–20 October 2006; Volume 50*, pp. 497–500.
68. Helmreich, R.L.; Foushee, C.; Benson, R.; Russini, W. Cockpit resource management: Exploring the attitude-performance linkage. *Aviat. Space Environ. Med.* **1986**, *57*, 1198–2000. [[PubMed](#)]
69. Röttger, S.; Vetter, S.; Kowalski, J.T. Ship management attitudes and their relation to behavior and performance. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2013**, *55*, 659–671. [[CrossRef](#)] [[PubMed](#)]
70. Bradley, B.H.; Postlethwaite, B.E.; Klotz, A.C.; Hamdani, M.R.; Brown, K.G. Reaping the benefits of task conflict in teams: The critical role of team psychological safety climate. *J. Appl. Psychol.* **2012**, *97*, 151–158. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).