A Research Agenda in Maritime Crew Resource Management

Michael Barnett, PhD; David Gatfield, MSc; Claire Pekcan, MSc

Abstract

This paper opens with a brief introduction to the development of Crew Resource Management (CRM) training in the international shipping industry, a concept that was first advanced through the use of simulators in maritime training colleges over 25 years ago. The paper charts the development of the shipping industry's approach to the preparation of bridge and engine room teams for normal and abnormal operations, and critiques the current training regime in resource management. Two case studies are presented to highlight some of the CRM issues raised by recent maritime casualties, and the paper then proceeds to set out a research agenda for exploring some of these issues. The paper provides an overview of three research initiatives: the first is to gain a better theoretical understanding of the nature of shared situational awareness and mental models in "real world" maritime operations. A second initiative is to identify a set of behavioural markers for assessing the non-technical skills of crisis management. The third initiative is to explore the role of organisational factors in safe operation, in recognition of the limitations of operator training as a panacea to prevent the re-occurrence of accidents.

The Development of Maritime CRM Training

The use of simulation in providing solutions to the problems of crisis management and the optimal use of crew resources has a long established pedigree in maritime training. The first simulators were introduced for radar training over thirty years ago. Training in the proper interpretation of radar information started as a result of a number of radar-assisted collisions in the 1950's, notably the collision between the passenger ship "Andrea Doria" and the "Stockholm". Those early simulators consisted of real radars, located in a set of cubicles, and fed with simulated signals. Individuals or teams could learn the skills of radar plotting under the guidance of an instructor working at a separate master console. Other navigational aids in the simulator were fairly basic and certainly did not include a visual scene.

Bridge simulators with a nocturnal visual scene made their appearance in the 1970's and allowed teams to conduct simulated passages in a realistic environment but with only a few lights available to indicate other vessels and shore lights. It was apparent from the casualty of the Very Large Crude Carrier (VLCC) "Metulla" in 1974, in which the vessel grounded in the Magellan Straits with two pilots and watch keepers present on the bridge, that bridge teams were not working effectively in supporting each other or the pilot. Simulator-based training courses were introduced primarily to train the skills of passage planning and the importance of the Master/Pilot relationship (Gyles and Salmon 1978). This training initiative developed into the Bridge Team Management (BTM) courses that are conducted today on many simulators world-wide and contain many of the elements to be found in CRM courses in other industries. Bridge Resource Management (BRM) courses are a more recent initiative, adapted directly from the aviation model, and are not always based on the use of simulators.

The 1980s saw the introduction of Engine Room simulators and towards the end of that decade, cargo operations simulators also became available. These types of simulator have primarily been used to train officers in the handling of operations, including fault finding and problem diagnosis, and increasingly to train teams in the skills of systems, resource and crisis management. Many types of simulator: bridge, engine and cargo control room, have tended to emphasise a physically realistic environment in which these exercises occur, although the use of PC-based simulators for training some tasks is increasingly widespread. In some parts of the world, simulators have been developed which have very high levels of physical fidelity, for example, multi-storey engine room mock-ups and bridge simulators including features such as 360 degrees day/night views, pitch and roll, and full vibration and noise effects.

Within other safety critical industries, and the military, the training and assessment of resource management skills is taking on a high level of importance as a way of ensuring that errors are effectively detected and managed (Flin & Martin, 2001; Cannon-Bowers & Salas, 1998; Brabazon & Conlin, 2000; Flin et al., 2000).

Evidence from some recent investigations into maritime accidents has shown severe shortcomings in the competence of some merchant marine engineering officers to manage both resources and crises (MAIB, 1994; MAIB, 1996; MAIB, 1999). Data from research undertaken by the UK Protection and Indemnity Club (UK P&I, 1997) indicates that human error directly accounted for 58% of all shipping incidents that led to major insurance claims. The United States Coastguard (1995) states that the human element was a root causal factor in 70% of all shipping incidents. Although not all of these incidents led to a crisis situation, all had that potential. Accepting that human error is inevitable, there is a need to understand the behaviours of effective error detection and management in order to ensure safe and efficient operations (Helmreich et al., 1998).

Although there is now a general acceptance of the core concepts for the non-technical or resource management skills required for competence in crisis management, there is also an acceptance that the behaviours associated with these skills are context specific. Helmreich et al. (1998) suggest that the optimal implementation of resource management skills is dependent upon the cultural context in which they are applied. If this is the case, then in order to effectively assess the application of resource management skills, assessment should, as far as possible, be undertaken within a cultural context close to that in which the skills would be applied. They discuss the influences of three cultures, professional, organisational and national, on the application of resource management skills in the cockpits of civilian airliners. They argue that resource management skills form a fundamental part of any error management philosophy and that these skills are highly applicable to any domain where teamwork and technology are required.

Within the maritime domain the only mandatory non-technical skills requirements are those of the International Maritime Organization's Seafarer's Training, Certification and Watchkeeping Code (International Maritime Organization, 1995). Table *A-V/2* of this code specifies the minimum standard of competence in crisis management and human behaviour skills for those senior officers who have responsibility for the safety of passengers in emergency situations. The competence assessment criteria detailed within the code are not based on specific overt behaviours, but rather on generalised statements of performance outputs, and as such are highly subjective and open to interpretation. Although the existence of these standards of competence indicates that the International Maritime Organization recognises the need for non-technical or resource management skills, both the standards and their assessment criteria are immature in comparison with the understanding of non-technical skills, and their assessment, within the civil aviation domain.

In summary, resource management training has become established in the curricula of maritime training establishments. Courses take a variety of forms and cover both deck and engine room disciplines. The courses are often simulator-based, but not always, and their syllabuses reflect CRM training in other industries. As can be seen from the history of this development, most major training initiatives have resulted from the lessons learnt from a succession of casualties.

However, the incidence of major marine casualties still continues and this raises the question of how effective training actually might be in improving safety performance. A recent review of the benefits of human factors training in aviation suggests that there are few published studies that demonstrate positive benefits in attitude or behaviour following the introduction of such training, and even those studies which do show benefit suffer from methodological weaknesses (Edkins 2002).

CRM Training: A Panacea For Operator Error?

The maritime training community often finds that the application of CRM style training is limited to a retro-active 'dose' of post incident remedial training. A collision or a grounding is likely to result in bridge teams being prescribed a course of 'treatment' in passage planning; an engine room fire or catastrophic failure is likely to result in engineering officers being prescribed a course of treatment in engine room management.

Quite rightly, the management of shipping companies feel that these potentially life threatening incidents need to be addressed. The human errors arising from poor judgement, poor situation awareness and procedural violations are unpacked to see what lessons can be learnt. Officers are sent on the courses in the hope that their erroneous behaviour can be un-learnt and replaced with more appropriate behaviour. In effect, training colleges are asked to fix 'problem employees'. The training colleges oblige with a week's course for the problem employees but it is unlikely that the course members will ever sail together as a team.

In our rush to fix the 'problem employee', we are all in danger of missing the point: different ships, different teams, different individuals, but the same sort of incidents keep occurring. Something more fundamental, more deep rooted than operator error is at fault. In the same way that having a documented safety management system does not make a company safe, having employees attend CRM courses does not in itself make a ship safe. Most company managers fail to ask why this is the case. Training is often seen as an end in itself and little effort is made to follow up the training by seeing how effective it has been on board the vessel itself. Little or no research is done to analyse whether solutions other than training are more appropriate and the training community unwittingly colludes in this self-deception by supplying yet more customised courses.

A recent casualty is now reviewed to illustrate the complex nature of the problem.

Case Study 1: The "Diamant" and "Northern Merchant" Collision: Professional Malpractice or Commercial Imperative?

On 6th January 2002, the Dover Strait, one of the busiest waterways in the world, was shrouded in thick fog. Visibility was less than 200 metres in places. Two ferries were crossing the Dover Strait at 0900 that day. The "Diamant" was coming from

Oostende heading for Dover. The “Northern Merchant” was heading to Dunkerque from Dover. Both vessels were travelling at close to normal cruising speed: “Diamant” a high-speed craft was travelling at 29 knots, and the “Northern Merchant”, a Ro-Ro ferry, was travelling at 21 knots. Were they to have continued their course and speed, the vessels’ paths would have taken them to within half a mile of one another. As it was, at just over a mile apart, the bridge teams started to question the assumptions they had made about each other’s probable course of action and started to implement course changes, but not speed changes, that would, they believed, put a greater distance between themselves. At 0952 they collided.

Analysis

- In open waters, the companies’ recommended distance to keep clear is 1 mile. It was not until the vessels were at the point of crossing this one-mile separation did they even start to question their assumptions about the actions of the other vessel and consider altering the course of their own vessel. Furthermore, these actions were all driven by what they saw on a radar screen, not what they could see from their bridge window, for they were navigating in thick fog.

  Employee fault: failure to maintain adequate distance = violation of procedures.

- The “Diamant”’s procedures stated that commercial considerations should not take precedence over good seamanship and in poor visibility, where there is any doubt, speed should be reduced. At 29 knots a ship travels one nautical mile in two minutes. Visibility was only a quarter of a mile. “Diamant” would have had to travel three-quarters of a mile into her safety separation zone before she could see the other vessel and all ambiguity would be removed. Thereafter, she would have approximately 30 seconds to take avoiding action.

  From the plot they were receiving on their radar, the bridge crew of “Diamant” thought that the “Northern Merchant” was on their starboard side and should therefore keep clear. The “Diamant”’s crew maintained course and speed and steamed on at 29 knots reducing their separation distance at a rate of a one-tenth of a mile every ten seconds. At about a quarter of a mile “Diamant”’s crew realised that the “Northern Merchant” had not kept clear and used their remaining 30 seconds to take avoiding action by significantly altering course – to port. Unfortunately, “Northern Merchant” was actually on her port side, and “Diamant”, having reduced their separation even further by turning to port, did not have enough time remaining to take avoiding action and steamed straight into “Northern Merchant”.

  Employee fault: failure to reduce speed = violation of procedures.

- The collision regulations state that when a vessel can only ‘see’ another vessel on its radar and a risk of collision exists, she shall take avoiding action in ample time.

  The crew of “Northern Merchant”, as do all professional merchant navy officers, know this rule very well. However, they were also in possession of another ‘informal’ rule: high-speed craft, due to their manoeuvrability, always keep clear. Unfortunately, this informal rule is inconsistently applied: “Diamant” did not keep clear because she thought she had right of way over “Northern Merchant”. “Northern Merchant” held on to the belief that “Diamant” would alter course until the very last moment, and then only made a small alteration of course, which was not picked up by “Diamant”. Given that “Northern Merchant” was travelling at

21 knots, one could argue that altering course at one mile (just under three minutes from a collision) was not in ample time.

Employee fault: poor professional judgement = violation of collision regulations.

On first reading the report, it is very easy to find fault with the actions of the bridge teams. Indeed, the Department of Transport investigation into the incident would concur that operator errors were the principal cause of the accident. Of the eighteen causes and contributions to the accident in the official report (MAIB, 2003; pp. 43-44) no less than eleven conclude that the actions of either the bridge team or the Master were a cause of the accident; in other words operator error. The recommendations that ensue are very much targeted at fixing these operator errors:

"The Maritime and Coastguard Agency is recommended to:

1. *Issue guidance to remind operators that Section II and III of the Steering and Sailing Rules of the Collision Regulations must be strictly complied with, acknowledging that vessels are not prevented from taking sufficiently early action, ahead of the point at which those sections come into effect.*
2. *Issue guidance on how operators should determine a safe speed and a close quarters situation in restricted visibility by:*
   - *Listing the factors to take into account, in addition to those prescribed in Rule 6 of the Collision Regulations; and*
   - *Defining the extent to which reliance can be placed on radar for detection of small vessels and other floating objects."* (MAIB, 2003; pp. 49)

Ostensibly, the Masters of both vessels were guilty of poor professional judgement, poor situation awareness, and violating procedures and collision regulations. Truly suitable candidates for some remedial training in bridge team management; or were they? If we subscribe to the 'problem employee' view, then we have to ask ourselves:

- Do these officers have no sense of responsibility for the safety of their passengers and crew?
- Are these officers by their nature risk seeking, enjoying the thrill of close-quarter situations?
- Is this why they failed to adhere to rules and practices with which they are very familiar?

It is difficult to imagine that any professional seafarers are cavalier about their responsibilities and reckless in the navigation of their vessels. So how do we account for their actions? The 'problem situation' view warrants some consideration.

At certain times of the year the Channel is often fog bound. Ships still continue to ply their trade. The routes between the ports of southeast England and the ports of northwest France are frenetically busy and competition for cargo and passengers is fierce. At the main ports on either side of this narrow waterway a 'first come, first served' system of berthing and unloading exists. This leads to time pressures on crews to beat the competition into port, as delayed berthing has a knock-on effect on customer service and thus commercial returns for the operator.

Given this situation, it is not difficult to see the extent of the commercial pressures on crews to maintain their turnaround times, even when it is foggy. They want to protect the profitability of their company because it is on this that their jobs and incomes depend.

The management of the "Diamant" had recognised the potential for problems to exist and had warned their Masters not to give precedence to commercial considerations when operating in poor visibility. Commendable, one might say. Yet how come the Master of "Diamant" was still operating the vessel at near normal speed? We have established that he was unlikely to be a maverick. We then have to ask:

- Was this a common occurrence?
- How many other Masters operated at close to normal speed in fog?
- How did the company monitor the compliance with this procedure?
- How often does the fleet cross the Channel in poor visibility in the same time that it takes to cross it in normal visibility?
- What were the rewards to the Master for compliance with this procedure and what were the punishments for non-compliance?

Without first establishing the answers to these questions it is not appropriate to embark on a course of employee fixing in the form of CRM. For, as soon as the 'fixed' employees return for duty, the same reinforcers of their old behaviour (there would be no reinforcers of the new behaviour) will still be in place. One can more or less guarantee that they would be operating their vessels at near normal speed in fog once again.

One important lesson to be drawn from this case study, therefore, is that it is incumbent upon operators to ensure that the words enshrined in their company procedures are backed up by actual management behaviour, in the form of monitoring undesirable behaviour of employees and taking appropriate actions to change the schedule of reinforcement.

Structure of the CRM Learning Environment: A Case of One Course Fits All?

As well as having to establish that CRM is the appropriate post incident response, it is also important that the structure and content of any CRM course, whether post-incident or preventative in nature, reflects the learning *needs* of the crewmembers. Of particular significance is the extent to which the learning environment is structured, requiring the learner to follow set procedures in response to events, and the extent to which it is unstructured, requiring the learner to implement their own solutions to the presenting problems. In this context, a distinction has to be made between emergency preparedness and crisis management.

An emergency can be defined as a situation outside normal operating parameters where corrective decisions and actions are based on documented procedures. In the maritime context, examples might be "Man overboard", steering gear failure or a report of a fire in a cabin. Emergency procedures can be trained both on board and at onshore training establishments.

A crisis differs from an emergency in that successful decisions and actions may not necessarily be based on documented procedures. Appropriate pre-defined responses may not exist, and even if they do, in practice they may have conflicting requirements. Those responsible for handling crises will have to think through the situation, and respond in creative and flexible ways.

This distinction between emergencies and crises has a significant impact on the training requirements for their management. Training in handling emergencies may simply be training in following pre-prescribed procedures and drills. Training in crisis management is likely to require a much more demanding approach to practise the skills required in these situations.

A recent casualty in British coastal waters is given to illustrate the need for training providers and purchasers of CRM training to be clear about whether they need to provide opportunities for their employees to learn crisis management or emergency preparedness.

Case Study 2: The "Green Lily" Grounding – An Unnecessary Tragedy?

On 18th November 1997, the 3,624 grt Bahamian registered vessel "Green Lily" sailed from Lerwick in the Shetland Islands with a cargo of frozen fish for the Ivory Coast. The weather on departure was bad with wind speeds increasing to severe gale force 9. The following morning, while hove to about 15 miles south-east of the island of Bressay in the Shetland Isles in storm force 10 winds, a sea water supply line fractured in the engine room. The engineers controlled the flooding and pumping out had begun when the main engine stopped. Unsuccessful attempts were made to restart the engine while the vessel drifted northwards towards Bressay. Shetland Coastguard was advised and three tugs, the Lerwick RNLI lifeboat and a coastguard helicopter prepared to proceed to the casualty.

Attempts were made by two of the tugs to secure a line and tow the "Green Lily" away from land but although initially successful, each line parted. The starboard anchor was released and the third tug attempted to snag the cable and pull her head to wind, but the cable parted. At this time, the lifeboat rescued five crewmen, including two injured, from the ship's deck. The ten remaining crew members were rescued by the Coastguard helicopter but the winchman, who had remained on the deck of the ship, was swept into the sea and lost. The "Green Lily" went aground and started to break up. The investigation by the Marine Accident Investigation Branch (MAIB), published in June 1999, advised the cause of the grounding was:

> *"the lack of propulsion and failure to restart the main engine to arrest the drift of the vessel towards the shore in the prevailing environmental conditions. Contributory causes included flooding of the engine room, failure to reset the mechanical over-speed trip, inadequate knowledge of the cooling water system, failure of the towage attempts and inadequate teamwork"* (MAIB, 1999; pp. 9)

Analysis

- An initial technical failure precipitated events and was compounded by a hostile environment and further technical problems and failures. The situation was escalating in severity. An emergency was becoming a crisis, but the actors in this tragedy did not have the benefit of hindsight to read the 'script'.

- The available emergency plans, which tended to be procedures based on single failures, were not applicable. The individuals involved were forced to fall back on their experience to cope with an increasingly complex and unpredictable set of circumstances.
- Initial diagnosis of the technical failure was incorrect and led to a faulty but persistent mental model of the situation. In this case, the chief and second engineers, together with the electrical engineer, failed to understand why the main engine stopped and were consequently unable to restart it. They believed that the main engine failure was due to the effect of the flooding, previously caused by the fracture of the sea suction pipe. The probable reason for the main engine stoppage was actually due to the mechanical over-speed trip either not being reset or reset incorrectly.
- Awareness of the overall situation by individuals was based on incomplete or inaccurate information. In this case, both the Master, based on his calculation of drift, and the engineers, were over optimistic in their belief that a tow would be available before the ship ran aground. Meanwhile, the skippers of the rescue craft had unexpressed reservations about various aspects of the operation including the appropriateness of some of the towing gear, the weather conditions and sea room, and the ability of the ship's crew to handle the towlines.
- Individuals and units were separated physically and several agencies were interacting through various forms of communication. In these circumstances, it was very difficult for the key players to communicate meaningfully and maintain a shared and agreed awareness of the rapidly changing situation.

The Research Agenda

In the year 2000, the Maritime Coastguard Agency (MCA), following a recommendation of the Marine Accident Investigation Branch (MAIB) in response to the loss of the "Green Lily", awarded a project to a research team at Warsash Maritime Centre. The remit of the project was to investigate the potential use of simulators for training in the handling of escalating emergencies. This project enabled the researchers to review current concepts and models in the field of crisis management across a range of safety critical industries and to conduct a survey of expert opinion on the optimal training and assessment regimes for handling escalating emergencies (Barnett et al 2002).

The project has also allowed the team to set out a research agenda for future work in the maritime context. The following sections describe three research initiatives in the field of maritime crisis management and CRM:

1. To develop a better theoretical understanding of the nature of shared situational awareness and mental models in "real world" maritime operations.
2. To identify a set of behavioural markers for assessing the non-technical skills of crisis management.

3. To explore the role of organisational factors in safe operation, in recognition of the limitations of operator training to prevent the re-occurrence of accidents.

1. Situational Awareness, Mental Models and the Paradox of RPD

Modern concepts for understanding decision-making have progressed from classic rational choice models to ones that try to reflect the way decisions are actually made in the real world. The most influential of these models is the naturalistic decision-making (NDM) model and has been defined as follows:

> *"The study of NDM asks how experienced people, working as individuals or groups in dynamic, uncertain, and often fast-paced environments, identify and assess their situation, make decisions and take actions whose consequences are meaningful to them and the larger organization in which they operate."* (Pruitt et al, 1997)

This definition reveals a number of characteristics of the situations in which NDM takes place:

- The situations in which decisions are made are uncertain, unpredictable and dangerous.
- Knowledge of the situation is incomplete, and constantly changing.
- The consequences of decisions and actions based on poor situational awareness are potentially catastrophic.
- Experienced people, not novices, generally conduct decision making in such situations.

Another important feature of NDM is that, unlike classical models of decision making, where the objective is to provide optimal decisions, the objective for real world decision makers is to arrive at actions based on decisions that will satisfy the immediate concerns of the situation, without those decisions necessarily having to be the best ones. There are a number of different models within an NDM approach to describe the process by which decisions are made. The dominant model is the Recognition-Primed Decision (RPD) model. Orasanu (1997) provides a comprehensive description of the process:

> *"Its basic principle is that experts use their knowledge to recognise a problem situation as an instance of a type, and then retrieve from their store of patterns in memory an appropriate response associated with that particular problem type. The response is evaluated for adequacy in the present context, and if it passes, it is adopted. If it is found wanting, either another interpretation of the situation is sought or a second level response is retrieved and evaluated."*

The RPD model works well to describe decision-making situations in the maritime context. But the model does have serious implications for the training of 'real world' decision-making skills.

In crisis situations, just when the expert needs to draw on a reliable repertoire, the situation is unpredictable and atypical, so no repertoire can be called upon. The crisis handler has to revert to a creative response i.e. they have to think their way through the novel situation. The primary justification for the direct training for crisis management is based in the belief that by exposing individuals or teams to a variety of potential crisis scenarios, their ‘patterns’ or mental models of situations will be enriched, thus enhancing their situational awareness techniques and their repertoires of decision making. The key to this approach is in the ‘richness’ of the mental models developed by the individual or team, but paradoxically, the problem is that if the training scenarios are too prescriptive, then the learned repertoires may be inappropriate to the real emergency encountered.

This repertoire driven process can lead to dangerous consequences when facing an unpredictable situation. On the one hand, the decision-maker may derive increasingly bizarre hypotheses to explain the available information cues – the “kaleidoscopic” effect; or the decision-maker may become fixated on one pattern, refusing to change repertoires in the face of obviously conflicting information – the “mind-set” problem as exhibited by the “Green Lily” engineers.

Decision-making is a skill. Like all skills, it may be honed through practice. By reducing cognitive load through practice, experts will be less stressed than novices in threatening situations. In addition to specific contextual skills, there is a set of more general decision making skills, - the metacognitive skills. The direct development of such generalised situation awareness skills might counteract the RPD paradox and the consequences of stress.

Major research issues to be addressed in the use of simulation for the training of crisis management include:

- To what extent will decision making skills, learned in a simulated environment, *transfer* to the real world?
- What level of *fidelity* has to be provided in the simulated environment to guarantee effectiveness of training transfer?
- How can the non-technical skills of crisis management be assessed most effectively? (see following section)

Two other research issues are of particular interest in the maritime context. The first is related to the sharing of situational awareness between members in a team and also between distributed teams. Both the “Diamant” and the “Green Lily” cases demonstrate difficulties in communicating mental models between teams on the same vessel and/or between separate agencies involved in a crisis situation. Video observations from our own simulator exercises suggest that team leaders can find it difficult to articulate their understanding of the situation to other team members. This difficulty is not limited to intra-team communication, but as the “Green Lily” case shows, can work at an inter-team level too. In addition, it is apparent that one team can easily become oblivious to the information needs of a separate team when under stress, for example, bridge and engine room teams habitually fail to update each other as a training scenario unfolds. Measuring the effectiveness of synchronous training and the characterisation of behavioural markers for distributed teams represent interesting challenges to the maritime training community.

The international shipping industry shares with the offshore industry a similar working environment in that multi-national, multi-cultural crews work and socialise together in an isolated environment for months on end. Cultural and linguistic effects on team working is a particularly challenging area of research. Our experience from simulator training suggests that different cultures do work together in noticeably different ways, for example, a UK/US team does display a more individualistic way of sharing situational awareness than those from a more "collective" culture (Hofstede, 1991). Questions that have yet to be addressed include:

What effects are produced by cultural factors and how may they be characterised? What is the impact on the overall safety performance of a team, especially in stressful situations, by placing individuals from one culture into a different culturally based team?

2. Behavioural Markers for Assessing Competence in Crisis Management

All safety critical organisations consider how they would manage a crisis situation and undertake some form of preparedness training. This training concentrates mostly on how to deal with an emergency, where a laid down procedure can be put into action. Few of these organisations take their training into the realms of a crisis situation, where there is no procedure to call upon, and where lateral thinking and rapid decision-making are required of their managers. Even fewer organisations try to assess their personnel's competence in managing a crisis. So how do safety critical organisations assess the competence of their crisis managers? How do they do this objectively, and what are the assessment criteria they use?

Of all the safety critical organisations, the military have taken crisis management training and assessment the furthest. This is done for a very good reason, as all combat situations are, by their very nature, crises. Confirming the experience of researchers in other domains, Tollcott (1992) states that the two primary components of military decision making are:

- situation assessment (what is happening); and
- action selection (what to do about it).

The first of these components requires crisis managers to generate hypotheses to account for the information that is being received. The second of these components requires the generation and evaluation of alternate actions. During a crisis these tasks have to be performed within a highly demanding decision environment. In certain circumstances this demanding decision environment may become too demanding for the crisis manager, and they may find themselves unable to cope. This is described by Salas et al. (1996) as a situation when:

> *"environmental demands evoke an appraisal process in which perceived demand exceeds resources and results in undesirable physiological, psychological behavioural or social outcomes."*

So it is important within any safety critical organisation to try and determine whether the personnel placed in the role of potential crisis manager will be able to cope when a crisis arises.

Military Behavioural Markers

Following their participation in a major US military research project, "Tactical Decision Making Under Stress", Cannon-Bowers and Salas (1998) proposed a set of twenty knowledge, skill and attitude requirements for teams to work effectively during crisis situations. If indeed there are so many requirements for an effective crisis management team, the assessment of competence in crisis management based upon these requirements is a daunting task. If assessment should be undertaken in an environment that closely resembles the real world situation, in order to ensure that, as far as possible, all of the required cues for decision-making are available; the capture of data to evaluate against assessment criteria relating to all of these requirements is a truly enormous task.

Through their use of war games, the military attempt this task. They use large numbers of assessors, dispersed throughout the war-gaming environment during an assessment exercise. After the assessment exercise, the assessors meet to discuss their observations during the exercise, and to evaluate the actions of the team against set assessment criteria. Examples of these criteria are given in Table 1.

Command and Control
Did the Command team quickly close up at HQ1.
Was a comprehensive set of check cards / aide memoirs provided and used.
Were smoke boundaries established and effectively maintained.
Was there a good flow of information into the control position at all stages.
Was the incident picture well kept.
Was an alternative control position considered.
Were hands piped to emergency stations in good time.

Table 1. Examples of Military Command and Control Assessment Criteria (Royal Navy, 2002)

These criteria are assessed as having been either ‘met’ or ‘not met’. A discussion is then held between assessors to give an overall assessment of how the team performed. Due to the severe time restraints imposed on the assessment process, subjective assessments are inevitable, because of the operational requirements of the military, and the sheer complexity of the war-gaming environment. However, because of the large number of assessors used, fair and effective assessments can be achieved through moderation. Anecdotal evidence from military personnel who have been assessed by such a process tends to confirm this fairness and effectiveness.

Civil Aviation Behavioural Markers

Within the civil aviation industry the training of crew resource management skills has been introduced as a way of improving safety performance. The civil aviation industry has recently been undertaking research into the possibility of assessing the non-technical skills of aircrew. Non-technical skills can be defined as those skills, in addition to technical skills, required for competence in crisis management. There are four main categories of resource management skills, or non-technical skills, being used within behavioural marker systems within the civil aviation industry:

- Co-operation
- Leadership and Management
- Situation Awareness

- Decision Making (Flin & Martin, 1998)

The European Union research project ‘Joint Aviation Requirements Translation and Elaboration of Legislation’ (JAR TEL Consortium, 2001) has evaluated the use of such a behavioural marker system for the assessment of resource management skills of commercial flight crews. The JAR TEL project concluded that their assessment framework: “is capable of proving itself a valid and reliable method for assessing non-technical skills.”

Although the ‘NOTECHS’ framework has moved the assessment of competence in crisis management, within the context of civil aviation, towards a more objective foundation, the experimental results of inter-rater reliability trials showed that in the more complex assessment scenarios there were some significantly divergent assessments.

The JAR TEL report states that there are some strongly held reservations, by some members of the aviation fraternity, about the very concept of the assessment of non-technical skills. One of the prime reservations being that: “it is felt that the criteria on which assessment is based are largely subjective and thus cannot easily be monitored for fairness and accuracy”.

Through the JAR TEL research project, a methodology for assessing the non-technical skills of aircrew, by observing individual overt behaviours, has been proposed. Some examples of the behavioural markers used in this assessment framework are shown in Table 2.

<table>
<tr><td colspan="2">Non-Technical Skill Category – Cooperation
Element – Consideration of others.
Consideration of others is about acceptation of others and understanding their personal condition.</td></tr>
<tr><td>Behavioural Markers indicating poor practice:</td><td>Behavioural Markers indicating good practice:</td></tr>
<tr><td>Ignores suggestions of other crewmembers.</td><td>Takes notice of the suggestions of other crewmembers even if s/he does not agree.</td></tr>
<tr><td>Does not take account of the condition of other crewmembers.</td><td>Takes condition of other crewmembers into account.</td></tr>
<tr><td>Shows no reaction to other crewmembers.</td><td>Gives personal feedback.</td></tr>
</table>

Table 2. Examples of Civil Aviation Non-Technical Skills Assessment Criteria (JAR TEL Consortium, 2001)

The cockpit environment is very different to that of a war-gaming environment, but the non-technical skills of co-operation, leadership and management, situational awareness and decision making, as metrics for assessing competence in crisis management, are common to both. As with the military assessment framework, the JAR TEL criteria are assessed as having either been passed or failed.

A major difference between the assessment of competence in crisis management within the military context, and within the civil aviation context, is that within the military context a team is assessed, whereas within the civil aviation context it is the assessment of an individual working within a team that is undertaken.

Maritime Industry Behavioural Markers

Through the STCW Code Table A-V/2 (International Maritime Organization, 1995), the International Maritime Organisation (IMO) has provided the competence specification of a minimum standard of competence in crisis management and human behaviour for those officers who have responsibilities for passengers. As with the civil aviation industry, these competencies relate to individuals working within a team. The required underpinning knowledge, understanding and proficiency, are stated for each competence, along with methods for demonstrating competence and criteria for assessing competence. Examples of these assessment criteria are shown in Table 3.

Competence – Establish and maintain effective communications.
Information from all available sources is obtained, evaluated and confirmed as quickly as possible and reviewed throughout the emergency.
Information given to individuals, emergency response teams and passengers is accurate, relevant and timely.
Information keeps passengers informed as to the nature of the emergency and the actions required of them.

Table 3. Examples of Merchant Navy assessment criteria from the *Specification of minimum standard of competence in crisis management and human behaviour*. (International Maritime Organization, 1995)

IMO does not differentiate between crises and emergencies, and the Table A-V/2 relates primarily to the management of emergencies, citing the use of procedures and actions in accordance with established plans as a criterion for evaluating competence.

From the examples above, it can be seen that safety critical organisations undertake the assessment of competence in crisis management in very different ways. Based upon observations within various safety critical organisations, Table 4 provides a summary of their use of crisis management assessment frameworks. Within the merchant marine context, the STCW assessment framework for crisis management and human behaviour is too open to interpretation to be effective.

Any framework for the assessment of competence in crisis management within the context of the merchant marine would not have the resources available to it that the military has. The civil aviation assessment framework for non-technical skills, although feasible to apply within the merchant marine context, has not yet been shown to be reliable in assessing competence in crisis management. However, at least the civil aviation industry is undertaking extensive research to address the issues of crisis management competence assessment. In contrast, very little research is being undertaken in this field with the merchant marine domain.

Table 4: Comparison of the Assessment Frameworks for Assessing Competence in Crisis Management Within Different Safety Critical Organisations

Context	Assessment Environment	Assessor(s)	Assessed	Assessment Criteria	Remarks
Military	High-fidelity War Game within real environment	Multiple Assessors Distributed throughout assessment environment	Team	Specific task orientated completion criteria.	Complexity of assessment environment leads to subjective interpretation. Fairness achieved through moderation.
Civil Aviation	High-fidelity simulator	Single	Individual working within team	Overt behavioural markers with examples given of good and poor practice.	Assessment framework difficult to use in complex scenarios leading to divergence of assessment.
Fire Service	High-fidelity simulator	Two	Individual working within team	Specific task orientated completion criteria.	Two assessors used to moderate subjectivity of assessment.
Offshore Oil/Gas	High-fidelity simulator and simulations onboard	Two	Individual working in a team	Specific task orientated completion criteria and some overt behavioural markers.	Two assessors used to moderate subjectivity of assessment
Merchant Marine	Simulations onboard and table top.	Single	Individual working within team	Prescriptive, but very open to subjective interpretation.	Assessment framework too open to interpretation by assessing authority.

Within the context of the civil aviation domain the behavioural marker systems for the assessment of non-technical skills are still the subject of research. Within the context of the commercial shipping domain a literature review has not found any such research being undertaken. Within the recommendations of the UK Maritime and Coastguard Agency's Research Project Report 467 (Habberley et al, 2001), there was a call for research to be undertaken to understand how behavioural markers may be used for the assessment of crisis management standards of competence within the commercial shipping domain.

> *"Crisis management standards of competence are ill defined and consequently so are their 'behavioural markers' by which the standard may be assessed. More research is needed in this area, particularly in assessing the team working competencies."*

Towards the Development of a Maritime Assessment Framework

A research programme is currently being undertaken at Warsash Maritime Centre that is intended to provide an understanding of how a behavioural marker system could be used to assess the competence in crisis management of merchant marine engineering officers within the context of a merchant vessel engine control room.

Behavioural markers that could be used to assess competence in crisis management within the context of a simulated merchant vessel's engine room control room are being determined. Experiments are being undertaken to investigate the efficacy of these behavioural markers to assess competence in crisis management, and it is intended that this research will then go on to show if these behavioural markers can be used as the basis for an objective competence assessment framework.

The aims of this research programme are:

1. To understand how behavioural markers can be used to objectively assess competence in crisis management of merchant marine engineering officers.
2. To develop and validate an assessment framework that utilises specific overt behavioural markers to facilitate the objective assessment of competence in crisis management of merchant marine engineering officers.
3. To provide the international maritime community with an understanding of how a behavioural marker system can be applied for the assessment of competence in crisis management of merchant marine engineering officers.

Data is being collected by observations undertaken during exercise scenarios within the Machinery Space Simulator at Warsash Maritime Centre. Patterns of interactions between engine room team leaders and other engine room team members are being compared to try and determine if there are any particular patterns that lead to the successful management of crises. The research is establishing and codifying the overt behavioural markers that define the successful patterns of interaction that are found. The research is restricted to overt behavioural markers. The research is not attempting to determine and analyse the cognitive processes underlying the patterns of interactions.

Groups of three experienced engineering officers are first given a familiarisation session within the Machinery Space Simulator. The groups are then given a briefing about the situation onboard the vessel, with each research participant given a team role. One participant takes on the role of the Chief Engineer, one as Second Engineer and one as Third Engineer. The simulator operator participates in the role of Extra Second Engineer, but only by radio. In this role the simulator operator can act as a facilitator for the not wholly scripted research exercise scenario. There is also a Mechanic available to work under the supervision of the Extra Second Engineer. Once simulator familiarisation has been achieved, team roles have been assigned, and the scenario briefing has been given, the research exercise is undertaken.

Observations are made with the researcher being an acknowledged observer. It is understood that there is the potential for reactivity effects using this method of observation. However, anecdotal evidence over many years from professional marine engineering officers using the Machinery Space Simulator at Warsash Maritime Centre has indicated that the closed circuit television and audio monitoring systems, used for observation, are unobtrusive, and do not cause any reactivity effects.

Behavioural Marker	**Characterisation**
Ratio of the degree of feedback control to the degree of predictive control.	Indication of the level of situational awareness.
The number of alternative hypotheses and actions communicated to team members.	An indication of teamwork and the building of a shared mental model.
Level of satisficing exhibited.	Considering only as many alternatives as needed to discover one that satisfies.
Communicating in a way that shares ones mental model.	Building, maintaining and refining the accuracy of the shared mental model of the team.
Relevance and timeliness of unsolicited information passed between team members.	A measure of the degree of congruence between the mental models held by individual team members.
Level of anticipation of other team members needs.	Indication of the level of situational awareness.
Level of anticipation of future actions and task requirements.	Indication of the level of situational awareness.
Focus is too much on the reduction of uncertainty.	Indication of a tendency towards analytical decision-making, and away from naturalistic decision-making.
Tendency to focus on one system at a time, thereby ignoring the dynamics of the complete system.	An indication of the lack of a situation overview.
Amount of sampling behaviour exhibited.	An indication of the updating of situational awareness and mental model.
Number of unfinished sentences.	A measure of uncertainty.
Delegation of work tasks.	A measure of the effective use of all team members, and the alleviation of overload.
Patterns of movement.	Interpretation of patterns of movement to determine degree of situation overview.

Table 5. Characterisation of observed behavioural markers.

The initiating event for the research exercise scenario is seawater ingress into the diesel oil service and settling tanks. The diesel oil service tank supplies both main diesel alternator engines. This initiating event is based upon an actual incident that led to the loss of the motor tanker Braer at Garths Ness, Shetland on 5th January 1993, with associated massive oil pollution. The significance of the observed behavioural markers in influencing the outcome of crisis development will be determined by statistical methods.

From the research exercises run to date a number of behavioural markers have been noted that may be shown to be significant. These are shown in Table 5. For those behavioural markers that are shown to be statistically significant, range statements will be proposed which codify the behaviour of the engine room team leader in relation to exemplars of behaviours associated with the effectiveness of the team leader's crisis management skills. Different assessors will then be asked to use these range statements to assess competence in crisis management. As these range statements are all nominal variables, the kappa statistic will be used as a measure of agreement between the assessors. The kappa value will be used to show the reliability of the proposed assessment framework.

3 Organisational Factors

The argument has been made earlier in this paper that the training and assessment of operators can only ever be part of the solution to reducing accidents. Organisational factors also play a significant part in accident causation. So what are the research issues in maritime operations, at an organisational level, which need addressing?

The analysis of human factors in accident causation is still relatively immature in the maritime world. Although databases held by the MAIB and other parties interested in the causal factors of accidents – e.g. insurers and classification societies – do include human error taxonomies, little analysis is undertaken to identify trends or patterns. Even less analysis has been attempted in assessing the significance or frequency of organisational factors such as the incidence of commercial pressure or the effects of organisational culture on accident causation.

The differences in organisational culture between shipping companies is a well known phenomenon, but there has been little work on understanding the effects of organisational culture on safe and efficient performance. Recent mergers in the industry have led to some interesting situations where the safety cultures of very different organisations have clashed. What is the impact of these factors on the safe performance of organisations?

In much the same way as we are striving to identify a set of behavioural markers to assess the competence of individuals, so there is a need to establish a set of organisational metrics to determine the competence of shipping companies to perform safely.

The maritime industry has concentrated on making the vessels safe and fit for purpose, an endeavour that has been largely overseen by classification societies. However, as it becomes increasingly apparent that technical and engineering solutions are necessary but not wholly sufficient to secure standards of merchant vessels, and as it becomes apparent that there is little mileage in continually focussing on operator error to prevent accidents reoccurring, the organisational or management issues can no longer be ignored. Indeed, in the maritime industry, the role of classification societies is expanding into operations. Classification societies are intervening in the standards of ship operations, a role that has been legitimised by the introduction of the IMO's International Safety Management Code (ISM Code).

Arguably, in the past, shipping enterprises were rendered incompetent to manage their operation safely by the amounts of prescription that were issued from all quarters: IMO, class, port, and flag state. These bodies developed the expertise and then enshrined it in rules and regulations. All the ship operator had to do was comply, 'unthinkingly'.

Goal setting legislation, of which the ISM Code is representative, turns this thinking on its head. The ship operator is now viewed as the expert and is, to a greater extent, left to his own devices to run his operation as he sees fit. The caveat is that he must control the risks arising out of his activities as far as reasonably practicable.

The ISM Code is fashioned on total quality management thinking. ISM guides the operator, pointing out the areas that he should address, requiring without stating as much, that the operator develop a safety management system. Inputs are controlled, processes are standardised, outputs are measured and the system is audited. As with other management systems, whether designed to engineer business excellence or effective health and safety management, ISM is a document-based system.

It is easy for an auditor, in such a document-based system, to see whether or not a shipping enterprise is doing what it says it will do. However, this paper trail does not necessarily allow an auditor to establish whether the system is effective.

Not enough is known about the parameters governing functioning and performance of management systems. There is little research evidence to indicate what makes a management system work or indeed what prevents it from working. Equally, not enough is known about the metrics that enable the status of a management system to be determined.

Organisations exist to do work that individuals cannot accomplish alone. As with the individuals that make up organisations, it is argued, observers can make inferences about the effectiveness with which that work is done. In the same way an individual can be deemed proficient or inept at their job by observing the way they behave, so can an organisation be deemed proficient or inept at transporting goods from A to B, making money for its shareholders, or whatever its principal mission is to achieve, by observing the way it 'behaves'.

The research conundrum is, first, to agree what constitutes organisational behaviour; second, in deciding which 'behaviours' are indicators of proficiency; and third, in designing methods that can measure these behaviours accurately.

Summary and Conclusions

As in similar safety-critical industries, the analysis of maritime accidents over the years has revealed shortcomings in the ability of operators to manage both resources and crises. CRM training has been seen increasingly as a fundamental part of the human error management philosophy. The International Maritime Organization recognises the need for non-technical or resource management skills, but both the standards of competence and their assessment criteria are immature in comparison with civil aviation. Although CRM training has become well established in the maritime curricula, as with civil aviation, there remains a question mark about how effective such training actually might be in improving safety performance. Analysis of recent casualties also suggest that CRM training, although important, may not be a panacea for operator error and that organisational factors must also be taken into account.

In setting an agenda for future maritime research in this area, we conclude that the following issues need to be addressed:

If the direct training of resource and crisis management skills is pursued, to what extent will such skills, learned in a simulated environment, transfer to the real world? What are the optimum training environments to ensure effective transfer? How can these non-technical skills be assessed most effectively, both at the level of the individual and at the level of the team? What behavioural markers, both at individual and team level, predict safe performance? In multi-national environments, how may cultural factors be characterised and what is the impact on overall safety performance of cultural differences? We know that organisational factors also play a significant part in accident causation but how can their significance, frequency and impact be established? Finally, what are the metrics that enable the status of an organisation's safety management system to be determined?

However, as with any scientific endeavour, it is not sufficient to just ask questions. There is much that is still not known about human behaviour in response to unexpected, unplanned and seemingly uncontrollable events. CRM training is a method that has been devised for preparing people to manage such events. The maritime community is to some extent playing 'catch up' with the research being carried out in the military and aviation arenas; and this is a privileged position. Maritime researchers are able to cogitate on the issues that their counterparts in other industries raise and it is their efforts that have inspired us to offer this maritime research agenda. It is offered, not as a guiding light for all now to follow, rather as a stimulus for debate. As a research community, interested in describing, predicting and ultimately, enhancing human performance, we need to make sure we are asking the *right* questions; questions that will lead us to conduct meaningful and fruitful research.

References

Barnett, M. L., Gatfield, D., and Habberley, J (2002) *Shipboard crisis management: A Case Study.* Proc Int. Conf. Human Factors in Ship design and Operation. pp 131-145 RINA, October 2002

Brabazon, P., & Conlin, H. (2000). *Assessing the safety of staffing arrangements for process operations in the chemical and allied industries.* Report for the Hazardous Installations Directorate of the UK Health and Safety Executive. London: HMSO

Cannon-Bowers, J.A., & Salas, E. (1998). *Making Decisions Under Stress: Implications for individual and team training.* Washington D.C.: American Psychological Association.

Edkins, G. D. (2002) *A review of the benefits of aviation human factors training.* Human Factors and Aerospace Safety. 2(3) pp 201-216 Ashgate Publishing.

Flin, R., & Martin, L. (1998). *Behavioural markers for crew resource management.* UK Civil Aviation Authority Paper 98005. London: Civil Aviation Authority.

Flin, R., O'Conner, P., Mearns, K., Gordon, R., & Whitaker, S. (2000). *Crew Resource Management Training for Offshore Operations.* Factoring the Human into Safety: Translating Research into Practice. Volume 3 of UK Health and Safety Executive report HSE OTO 2000 063. London: HSE.

Flin, R., & Martin, L. (2001). *Behavioral Markers for Crew Resource Management: A Review of Current Practice. The International Journal of Aviation Psychology*, 11(1), 95-118.

Gyles J.L. & Salmon, D. (1978) *Experience of Bridge team Training using the Warsash Ship Simulator.* Proc First Int Conf. on Marine Simulation MARSIM (1978) pp 1-26 Nautical Institute.

Habberley, J., Barnett, M., Gatfield, D., Musselwhite, C. & McNeil, G. (2001). *Simulator Training for Handling Escalating Emergencies*. United Kingdom Maritime and Coastguard Agency Research Project 467. Southampton: The Maritime and Coastguard Agency.

Helmreich, R., Wilhelm, J., Klinect, J., & Merritt, A. (1998). *Culture, Error and Crew Resource Management*. Austin: University of Texas at Austin.

Hofstede, G. (1991) *Cultures and organisations: Software of the Mind.* London: McGraw-Hill.

International Maritime Organization (1995). *Seafarer's Training, Certification and Watchkeeping Code (STCW Code)*. London: IMO.

JAR TEL Consortium (2001). *Joint Aviation Requirements: Translation and Elaboration of Legislation.* Final Report, Contract n° AI-97-SC.2037. European Commission, DG TREN.

Marine Accident Investigation Branch (1994). *Report of the Chief Inspector of Marine Accidents into the engine failure and subsequent grounding of the Motor Tanker Braer.* Southampton: MAIB.

Marine Accident Investigation Branch (1996). *Report of the Investigation into the Power Failure on Canberra.* Southampton: MAIB.

Marine Accident Investigation Branch (1999). Marine Accident Report 5/99. *Report of the Inspector's Inquiry into the loss of MV Green Lily*. Southampton: MAIB.

Marine Accident Investigation Branch (2003). Marine Accident Report 10/03. *Report on the Investigation of the collision between Diamant/Northern Merchant*. Southampton: MAIB.

Orasanu, J.M (1997*). Stress and naturalistic decision making : Strengthening the weak links.* In 'Decision making under stress: emerging themes and applications' (eds: R Flin, E Salas, M Strub and L Martin) Ashgate.

Pruitt, J S, Cannon-Bowers, J A, and Salas E. (1997*) In search of naturalistic decisions.* In 'Decision making under stress: emerging themes and applications' (eds: R Flin, E Salas, M Strub and L Martin) Ashgate.

Royal Navy (2002). *Basic Operational Sea Training Assessment Form.* Plymouth: HMS Drake. Flag Officer Sea Training.

Salas, E., Driskell, J., & Hughes, S. (1996) *Introduction: the study of stress and human performance.* In: J. Driskell & E. Salas (Eds) Stress and human performance. Mahwah, N.J.: Erlbaum, 1-45

Tollcott, M.A. (1992). *Understanding and aiding military decisions.* In: S.C. Collyer (Ed) The 27th International Applied Military Psychology Symposium: A focus on decision making research. London: Office of Naval Research, ONR Europe Report 92-4-W, 33-48

UK P&I Club (1997). *Analysis of major claims – ten-year trends in maritime risk.* London: Thomas Miller P&I Ltd.

United States Coastguard (1995). *Prevention Through People,* Quality Action Team Report. Washington D.C.: USCG.