Managing Escalating Situations at Sea: Testing and improving the value of ship-bridge simulation for maritime crisis management

Sidney Dekker, Nicklas Dahlström, Roel van Winsen and Johan Bergström

Lund University School of Aviation
Managing escalating situations at sea:
Testing and increasing the value of ship-bridge simulation for maritime crisis management

Hantering av upptrappande situationer:
Undersökning och försämring av värdet av bryggsimulering för hantering av maritima krissituationer

Interim status report

Sidney Dekker, Nicklas Dahlstrom, Roel van Winsen & Johan Bergström

Summary

Since the report in May 2007 the project has completed the planned data collections and collaboration with maritime educational organizations. In addition, contacts with the Swedish Rescue Services have resulted in data collections with Fire safety engineers. The project has clearly raised questions of great importance in regards to training for unexpected and escalating events in the maritime communities as well as in other organizations. The article presented in the last report, which was then submitted to an international journal, was accepted on conditions of revisions which could not be accepted by the project. This article has been reworked and submitted to another international journal. Another article has been submitted to an international journal and still another will be completed and submitted before the next report. The research project has been presented at national and international conferences seminars and meetings. The project has received great interest and there are plans for further cooperation with the maritime organizations as well as with the Swedish Rescue Services. Also, recently the project has been approached with requests from the medical sector regarding participation in the training/data collections and further cooperation. This report outlines the current overall status of the project as well as details regarding results, conclusions and planned work up to this point of the project.

Activities performed in the project

Data collections

Data collection with maritime students at Chalmers department of Shipping and Marine Technology, 26-27 April and 3-4, 24-25 May, Gothenburg (3 groups)

The cooperation with Margareta Lützhöft at the department of Shipping and Marine Technology at Chalmers University of Technology and her team was completed with a third and final data collection in late May. The data from the participation of three groups of maritime students, including senior students from the master mariner and naval engineering program, will be presented in a coming article. The fact that almost all of the students had some or extensive seafaring experience is a fact that will be of interest in this analysis.

Overall the cooperation with Chalmers was very positive for the project as it provided opportunity to discuss the problems investigated in the project with researchers and
participants with knowledge and experience of the maritime domain. Overall the groups were positive in regards to the M/S Antwerpen simulation and training concept and found it relevant for their future professional roles. The comments of the third group supported the previous group’s statements that the training provides an appreciated “learning by doing” – approach not found in other courses. It seemed like the participating students had enough knowledge to provide further input on potential improvements of the realism of the simulation, but not experience enough to be disturbed by such inconsistencies during the training.

The students were credited for the training as a part of their Crowd and Crisis Management course. This and the overall positive impression the training made has provided a potential for further cooperation in regards to the M/S Antwerpen training as a part of regular training for ship crews. Contacts will be taken to see if the discussion on a presentation of a demo of M/S Antwerpen within the faculty at Chalmers, and its potential use as a replacement of their current scenario used for Crowd and Crisis management training, can be planned within the frame of this project.

Demo of M/S Antwerpen and data collection at Crowd and Crisis Management course for Stena Line, 30 May, Gothenburg (1 group)

After our data collections at Chalmers the interest from the Human Factors Research Group in M/S Antwerpen was so great that they wanted to be able to offer a demo version to one of their most important regular customers for emergency management crew training, Stena Line. This demo was made to a group of operational staff, most of them with many years of experience in their professional roles, from the shipping company as part of a scheduled training package received at Chalmers.

This group proved was similar in their performance to the one with experienced seafarers experienced at World Maritime University (WMU). Although the cultural diversity was far less for this group, the professional diversity of this group was greater since other categories of staff than nautical (catering, sales etc.) were represented. In their performance in the simulation they did however rely heavily on procedural solutions and showed little flexibility in their approach to manage unexpected situations. There was an obvious lack of understanding of different roles and the need for cooperation outside of roles as complexity and escalation of a situation increases.

The group found the training interesting and relevant and could see a potential for its use in regular training for ship crews. The group also provided some valuable input in regards to some details in the simulation that could be improved. They did explicitly state they would also want the opportunity to perform this training on a non-domain specific scenario. Like with the demo for BRM workshop leaders, it is difficult to use the demo version to collect data that is comparable with the regular M/S Antwerpen concept. (Also, because the group was large and to let everyone participate it was necessary for some people to share roles and relive each other during the simulation, a practice is outside of the normal procedures for the training.) It was however clear that in many aspects the performance of this group was similar to the group at WMU (and also to the participants at the previous demo for BRM Workshop leaders at SAS Flight Academy) and supportive of conclusions drawn from those sessions.
Data collection with Fire Safety Engineers, 16-17 August and 5-6, 10, 13-14 September, Revingehed (2 groups)

Initial contacts with the Swedish Rescue Services School at Revingehed to get access to M/S Antwerpen participants were met with great interest. This developed into a collaboration which provided access to Fire safety engineers who were studying a post-graduate practical course in emergency management. At the time of the data collections the Fire Safety Engineers were more than half-way through their one year education “Rescue Services Training for Fire Safety Engineers”. A class of 24 students was divided into four groups and two of the groups received the M/S Antwerpen training.

This was followed up with observation and data collection during the following simulated exercises that were part of their education. These “staff exercises” were focused on management of crises. The first scenario included a plane crash and the second an accident with a trailer filled with chemicals in an urban environment. All four groups were observed and videotaped during these scenarios and questionnaires were collected from both participants and instructors.

The M/S Antwerpen training was highly appreciated by the Fire safety engineers who received it. They considered it as relevant and valuable practical training which they would have liked to go through earlier in their practical training program and had more of. The two groups who did not receive training expressed concerns that they also wanted the training and was satisfied to learn that they would get it later. These follow-up data collections will be performed in December and may provide interesting data since the participants should by then have completed their staff training.

Data collection with Flight Instructors at Lund University School of Aviation, 1-2 November, Ljungbyhed (1 group)

Initially this project collected data with student pilots at Lund University School of Aviation. Since then it has been our ambition to do a data collection with the Flight Instructors to see if this would provide us with any different data in comparison with student pilots or with other experienced professionals groups. Due to the recent nature of this data collection we have not had time to review and analyze the data yet. Results from this data collection will be included in the next report from the project.

Planned data collections

The main focus for the project will currently be on analyzing the data for the collections performed during this year. There is however still a need to collect more data to be able to support conclusions being considered as well as to test new hypotheses suggested by observations during the simulation. One such observation is the normally “informal nature” of work in a group of participating maritime or aviation students and how this informal structure becomes inadequate as escalation starts to make a situation unmanageable. There seems then to be a difficulty for these groups to “upgrade” to a formal structure of work in order to manage the workload effectively. This seems to be a problem at the other end of the spectrum in regards to the steep hierarchical structure seen in the WMU-group.
Another two groups of Fire Safety Engineers will however go through the training program in December to conclude the training for all of the participants in this part of the project. This will also provide further opportunities for interestingly framed data collection since these participants during the autumn have received training in crisis management, in particular regarding effective procedures for working in a staff of a crisis management group.

There have been requests for further cooperation from the World Maritime University (WMU), as discussed at the conference MARAD at WMU (see listing of presentations). And this potential will be followed up on in the project. The cultural diversity found at WMU provides a potential for interesting comparisons with the more homogenous group found in other participating organizations.

Meetings with Professor Stefan Strohschneider, 18-19 and 29 June

The M/S Antwerpen simulation was originally developed at the University of Bamberg by Professor Stefan Strohschneider and the programmer Jürgen Gerdes in a project financed by the University Hospital of Aachen (after a less than adequate response to a large fire on the hospital in spite of that plans and procedures for such a situation were available and trained). In a meeting in June at his current workplace, the University of Jena, Professor Strohschneider was updated on the status of the project and its future was discussed. Plans for the project as well as potential publications and further collaboration on issues of simulation for training of unexpected and escalating events were discussed at the meeting.

The commercial rights for the M/S Antwerpen simulation are today held by the University Hospital of Aachen, something that is an obstacle to further development of it. At a subsequent meeting in July in Bamberg a representative of this project met with Professor Strohschneider and Jürgen Gerdes to discuss the potential to “build a new ship” in order to implement improvements that has come out of the experiences of this project. In addition, the potential to build a viable marketable training concept based on the M/S Antwerpen and the experiences of this project, e.g. including improved documentation and facilitator training, was explored. Further contacts on these issues are planned during next year. Although these are currently only ideas they indicate that the project has advanced the simulation and training concept to a level where further steps of its use can be planned.

Meeting with development group of the Swedish Rescue Services, 13 September

At a meeting with the development group for leading and coordination within the Swedish Rescue Services there was great interest in the project expressed, specifically in aspects of simulation and training, was expressed. This indicates that the focus on general principles of management of escalating situations and how to train these is of great interest not only from transportation industry. Further contacts with the Swedish Rescue Services may open up opportunities for data collections with new groups, where experienced members of the Swedish Rescue Services would be an interesting alternative.
Meeting with Kalmar Maritime Academy, 9 November

Although no data collections have been performed up to now in collaboration with Kalmar Maritime Academy contacts have continued and at a recent meeting representatives at Kalmar were updated on the project and its progress. There are still intentions from both parties to try to perform data collections at Kalmar Maritime Academy.

Experiences of running the M/S Antwerpen simulation

Simulation procedures and documentation

The procedures for running the simulation are well established in the project. This includes emergency planning in the sense that in one session there was a computer malfunction required the simulation to be restarted. The procedure for this had been worked out previously and it all worked well, with the participants resuming the simulation at the same point of time where it was interrupted after a fifteen minute coffee break. There is however still a need to further document the experiences made into procedures and advice for future facilitators. This is a task that there will be focus on during the coming last year of the project.

One area where there is potential for further improvement is documentation of explanations and motivations for the development of certain scenario events. One example of this is the failure of the sprinkler system to take out a large fire, where contacts with expertise in fire safety has provided a credible scenario for this as well as clues on this provided to participants during the run of the simulations. This has improved the credibility of the simulation and almost eradicated participant questions or frustration on issues of the spread of a fire after the simulation. Similar improvements can be made regarding other aspects of the simulation.

Participant information

The participant information is due to numerous corrections and revisions reasonably consistent and correct. The terminology of the material is now better aligned with correct maritime wording for positions. Also, with the experience from the sessions with fire engineers and the inclusion of a project team member with fire engineer competence the descriptions of the fire fighting equipment and associated systems have been improved. Although there are further items for correction discovered in recent data collections they remain at a small number as compared with previous sessions. Further cooperation with maritime organizations should make it possible to improve the participant information in regards to use with maritime groups; with non-maritime groups however there does not seem to be a need of further improvement

There are however in the simulation itself a number of errors and inconsistencies which can only be addressed with reprogramming of the simulation. (An example is that the abbreviation “sm” is used in the simulation for “sea miles”, instead of the correct “nm” for “nautical miles”.) Due to the issue of property rights to the M/S Antwerpen simulation (as outlined previously in this report under the heading “Meetings with Professor Stefan Strohschneider”) corrections of this kind, which requires reprogramming of the original code of the simulation, will not be made in the frame of this project.
Facilitator training

Although the requirement of having facilitators available for data collections has been obvious and there has been an additional facilitator trained the project still has only two facilitators easily available. Although the training of new facilitators have been expediently carried out in an informal manner this has highlighted the need to develop a formal and documented training for facilitators of the M/S Antwerpen. Currently there is no such training available and the expertise of running the simulation resides with Professor Strohschneider and the most experienced members of this project group.

As ideas of facilitator training has been sketched it has become clear that this would be a necessity if the mid-fidelity training of this project is to be widely used as a training tool for various groups that may need to train for unexpected and escalating crisis scenarios. To this end it would probably be prudent to consider at least to training concepts, one for participants who have gown through the M/S Antwerpen training and one for novices.

Revision and addition of simulation events

The new scenario events introduced during spring have successfully added to the credibility and complexity of the simulation. The new events have also been presented and discussed with professor Strohschneider and appreciated as an innovative and constructive step in the development of the simulation. With these new events there should not be room for further additions of new events.

One example of this is the addition in the second scenario of an event where some caged birds are discovered in a cabin. This is a minor and non-important event when considered properly. Since it is introduced during an idle period and followed by more consequential events it does however have very interesting effects on the workload and priorities of the participating group. Even when considering that being careful with the introduction of new events is of great importance we can conclude that the addition of events that has been performed in the project during this year has been successful.

Demonstration version of the simulation

As the demo version has been used for another maritime group we can conclude that it provides an option to present the M/S Antwerpen simulation in a limited time frame. When to use the demo version’s two-hour simulation trip (instead of two three-hour trips) and abbreviated presentations is however a difficult choice. Although it may raise interest in the project and the simulation it can only present a fragment of the training as there is limited opportunity for a participating group to reflect on its performance and no opportunity to provide relevant theory, facilitator support or to improve performance in a second session.

Since the demo version has proved successful in presenting the training concept and research program it should be an option to use it in the project. It should however never be presented as a training opportunity. The demo program is to be used to gain industry acceptance and opportunities for data collection. In addition it may provide access to groups with expertise that could support improvement of the simulation and training concept. From the project we will however stress the demo version removes the contextual and pedagogical content from the training and can only be used for a training technology oriented demonstration.
Methodological issues

Data collection

The previous report brought up the problem of collecting unbiased data from participants during the debriefing. In response to this a questionnaire for the participants was constructed. It has been used for all groups since the last report to collect data on the experiences (e.g. regarding, goals, information sharing, workload, decision making, roles) connected to the simulation and training. After evaluating this procedure we have become convinced that we should return to use the debriefing as an autonomous group process guided by a debriefing instruction and followed up with a facilitator-supported discussion.

The questionnaire seems to prompt too much of “correct” answers that previous groups of participants have managed to conclude themselves without any guidance other than a general debriefing instruction. This may lower the commitment to try to implement the new ideas about how to work in the second scenario. Also the answers to the questions have indicated that immediately after the simulation it is difficult for participants to evaluate their own behavior. In the subsequent facilitator led discussions participants have routinely changed their view of their own and the group’s performance as the group review their performance together. Thus, we are now aiming at returning to the previous procedure as mentioned above.

The current debriefing instruction does however need revision in the light of the experiences made with the M/S Antwerpen simulation and the questionnaire. Until the next report from the project we expect to have produced a new debriefing instruction as well as facilitator instructions on how to use it to ensure optimal effect of the training on the group.

Regarding technical matters sound recording is still a problem with unreliable microphones occasionally creating a less than adequate recording. New microphones will be tested in the coming data collections to improve the reliability and quality of the sound recordings.

Papers and presentations of the project since the last report

Articles

In the last report we presented an article submitted to the International Journal of Professional Aviation Training & Testing Research. The response from the journal was that the article would be accepted with revisions. The suggested revisions were however judged as unacceptable by the project group and it was decided to aim for a journal where more qualitative results and general discussions as well as conclusions could be presented. Since this decision the article has been reworked and submitted to Theoretical Issues in Ergonomics Science, a journal where more general questions regarding simulation and training can be presented with greater impact. In addition another article about the project have been submitted to the journal Disaster Prevention and Management


Presentations

The research project has been presented to different audiences at the events below, in most cases as main focus of the presentation and in some cases as central parts of presentations on safety, Human Factors, training and simulation.

*Can we still learn from aviation?*
Annual meeting for the Society in Europe for Simulation Applied to Medicine, 18-20 June, Copenhagen, Denmark.

*Human Error – Models and Management*
International Summer School on Aviation Psychology, arranged by University of Graz, 1-6 July, Graz, Austria.

*Safety Culture in Aviation*

*Cooperate, lead and decide*
Meeting of development group for leading and coordination, 13 September, Revingehed

*Training of escalating crises with simulation*
Seminar on Learning for Rescue Services, 2 October, Karlstad

*Safety and error*
Crew Resource Management course for the Swedish Armed Forces, 12 October, Ljungbyhed

*Human error and safety*
Patient Safety Course, University Hospital of Malmö, 15 October, Malmö

*How should we get people to do what they should do?*
Course for managers with responsibility for explosive materials, Competence Centre for Energetic Materials, 15 October, Örebro

*Safety and error - based on work with aviation safety*
Cambrex (Chemical Industry) Annual Safety Day, 19 October, Karlskoga

*Errors in understanding of human error*
Seminar for instructors for nuclear operators, Nuclear Safety and Education, 6 November, Studsvik

*Parallels between Aviation and Patient Safety*
Regional Conference on Patient Safety, Lund University Hospital, 8 November, Lund
Deliverables

M/S Antwerpen simulation

In the period since the last project report the M/S Antwerpen mid-fidelity simulation and the training concept have been further improved regarding documentation for facilitators and for the simulation, establishment and evaluation new scenario events and some developments regarding the theoretical training. Also, these developments of the simulation and training have been presented and discussed with Professor Stefan Strohschneider to anchor them in the frame of the original research that initiated the construction on the simulation and training concept.

In addition two articles have been submitted and a third will be finalized and submitted before the next report. Also, numerous presentations to communicate the results and conclusions of the research project have been given.

Planned articles

It is expected that at least another article will be submitted for publication before the next report is due. There is also potential for further articles as indicated by the outline of ideas below:

*General competence in escalating situations*
This article will present the results from the data collections with the Swedish Rescue Services along with a theoretical base for the concept and relevance of general competencies and how they can be effectively trained.

*A study of cooperative behavior using mid-fidelity simulation*
This will be based on observations of student pilots during Multi Crew Cooperation – training. One group will have received the M/S Antwerpen training while another one has not and differences in performance will be presented and discussed. An instrument for categorizing behaviors is used in the study of the behaviors.

Appendices

**Appendix A.** Preliminary findings from data collection with Fire safety engineers

**Appendix B.** Article submitted to Theoretical Issues in Ergonomics Science

**Appendix C.** Article submitted to Disaster Prevention and Management

**Appendix D.** Preliminary version of article
Appendix A. Preliminary findings from data collection with Fire safety engineers

Preliminary findings from data collection with Fire safety engineers

Since this data collection was performed very recently there have been no analyses of the data performed yet. The aim here is to mediate some of the direct observations made by the facilitators during these sessions.

As with previous groups, a lack of integration of contextual aspects of the simulated situation contributed heavily to the groups performances during the first scenario of the simulation. It is clearly stated that the training is part of a research project funded by the Swedish Emergency Management Agency, that the training is focused on crisis and disaster management and that a simulation will be used. In addition to this the initial presentation outlines the often observed circumstances of a crisis and stresses the limitations of procedural knowledge to resolve crisis situations. In the initial presentation it is also emphasized in text and by the facilitator that that no assumptions on the conditions of or situation on the vessel should be made.

These contextual factors, implicit and explicit, did not prompt these groups of Fire safety engineers to formulate goals or even questions regarding fundamental safety conditions of the ship. Safety equipment was not checked, procedures for risk scenarios and evacuation were not explored and practically no proactive actions were taken during the simulation. Even if this is “normal” behavior in the first scenario of the M/S Antwerpen it was still somewhat surprising given the educational background of these groups.

Both groups did however improve their performance significantly to the second scenario. Even though this is prompted by the training setup and normal for group participating in the training the groups performed very well and used the potential they had in the form of previous education in fire safety and emergency management. In this sense their performance was similar to that of the maritime students, who also managed to get a significant leverage out of their domain skills in the second scenario.

Among the most interesting individual observations was that of one of the groups established a system for coordinated judgment and action in the form of a color code for the status of the vessel during the simulation. This was simply an agreement to let a color (green, yellow and red) on the whiteboard represent the status. As simple as this may appear it seemed to provide a helpful feature for the work of the group. The discussions of what color that should represent the current situation seemed to fill the function of focusing the group on potential threats and were naturally followed up with discussion on how to manage these threats. Also, by considering the severity of threats before taking action this group managed to not get drowned in minor problems (as has often been observed in other groups).

During the following regular simulated exercises there are indications from both observations and from instructors that the groups who received M/S Antwerpen training did perform better than the groups who did not. The data from both the M/S Antwerpen training and from the all of the groups in the subsequent emergency management simulations integrated in their will be further analyzed and complemented with data from the two groups who are planned to go through M/S Antwerpen training later.
Fidelity and Validity of Simulator Training

Nicklas Dahlström1, Sidney Dekker1, Roel van Winsen1 and James M. Nyce1,2

1Lund University School of Aviation, Drottningvagen 5, 260 70 Ljungbyhed, Sweden
2Department of Anthropology, Ball State University, 2000 W. University Ave., Muncie, IN 47306, USA

Through a case study, this article explores a number of theoretical issues related to the often taken-for-granted relationship between simulator fidelity and the quality and transferability of training in complex, dynamic safety-critical settings. We present a counterexample based on mid-fidelity simulation and test the assumed coincidence of fidelity and validity, that is, the equation of constructed photorealism (built to mimic reality) with effective development of the competence operators require to manage situations that involve underspecified problems, time pressure constraints and complex group interaction. We conclude that such competence development cannot rely only on highly context-specific (photorealistic) environments. Further it will be argued that lower-fidelity simulation, when appropriately designed, can provide competence development with pedagogical and economic advantages.

Keywords: simulation, fidelity, training

1. Introduction

Much of training in one operational world today—aviation—focuses primarily on “technical skills” (see Dahlström, Dekker & Nählinder 2006), i.e. the build up of an inventory of proceduralised interaction with the technical systems of an aircraft. From the start, this training is quite context-specific: it is set in, and tightly anchored to, the local technical environment (of an aircraft cockpit or a simulator with as high a fidelity as economically feasible) in which problem-solving activities are to be carried out. This is analogous to other domains that assert the primacy of learning-by-doing (for example surgery, see Bosk 2003). This means that generic competencies like problem-solving, communication, coordination or management of unanticipated and escalating situations are largely to arise, and emerge only during the exercise of context-specific work. Such exercise is valorised by the increasing technical sophistication of simulated training environments, in worlds ranging from aviation to shipping to healthcare. This article explores a number of theoretical issues related to the often taken-for-granted relationship between simulator fidelity and the quality and transferability of training in complex, dynamic safety-critical settings. We present a counterexample based on mid-fidelity simulation and test the assumed coincidence of fidelity and validity, that is, the equation of constructed photorealism (built to mimic reality) with effective development of the competence operators require to manage situations that involve underspecified problems, time pressure and complex group interaction. Can such competence development rely on highly context-specific (photorealistic) environments alone? Can lower-fidelity simulation, when appropriately designed, provide competence development with pedagogical and economic advantages?

1.1 Fundamental surprises

Operational activities in any industry contain situations whose subtle and infinite variations may mismatch the exact circumstances of training. This includes surprises, i.e. situations that

1 Corresponding author. Email: nicklas.dahlstrom@tfhs.lu.se
fall outside of the procedures constructed for normal and emergency operations. On these occasions operators must be able to apply skills and knowledge that no training department was able to foresee or deliver. This leaves a residue of potential problems that crews are not prepared for (i.e., they are not in their inventory; see Dismukes, Berman & Loukopoulos 2007). The same is for other industries and operators as well. Formal mechanisms of safety regulation (through e.g., design requirements, policies, procedures, training programs, line checks) will always fall short in foreseeing and meeting the shifting demands posed by uncertainty, limited resources and multiple conflicting goals. For this residue we have to count on crews’ generic competencies.

These surprises at the margins of an otherwise very safe system stem from limits in the industry’s knowledge, or, more often, limits on its ability to put together diverse pieces of knowledge, as well as from limits on the understanding of operational environments (Lanir 2004). In other words, the knowledge base for creating safety in complex systems is inherently imperfect (Rochlin 1999). Often the problem is not that the industry lacks the data. After all, the electronic footprint left by any commercial flight today is huge. The problem is that this accumulation of noise and signals can muddle both the perception and conception of “risk” (Amalberti 2001, Dekker 2005). Pockets of expertise that may have predicted what could go wrong often exist in some corner of the industry long before any accident.

An example is the Swissair 111 accident, where the presence of smoke in the cockpit was responded to promptly by the crew by adhering to the relevant checklist. However, while following the established and trained procedures fire engulfed the aircraft (Transportation Safety Board of Canada 2003). This accident revealed gaps not only in the industry’s knowledge of electronic wiring and insulation (and its inability to stem the push for more and more wiring through the bodies of aircraft). It also reminded the industry of the immense difficulty crews face in making trade-offs on adapting plans and procedures under duress and uncertainty, and of its shortcomings in preparing them for incidents like these (Dekker 2001).

More recently, the accident of Pinnacle Airlines flight 3701 exposed limits in knowledge of how the absence of passengers during repositioning flights can erode crew operational margins, and further revealed lacks in crew knowledge of high-altitude climbs, stall recognition and recovery from double-engine failures (National Transportation Safety Board 2007). The type of engine on the Pinnacle type aircraft had a history of problems with in-flight restarts during flight tests. Problems with wiring and chafing had also been noted prior to this incident. But few or no operational crews would have been aware of any of this in part because of structural industry arrangements that regulate who gets or needs to know what, and in what depth.

1.2 Resilience and the limits of expertise

As a result, some crews will, at some point or another, be left to “fend for themselves” at the edges of an extremely safe industry. It is at these edges that the skills trained for meeting standard threats need transpositioning to counter threats no one has not foreseen. The flight of United Airlines 232 is an extreme example (National Transportation Safety Board 1990). The DC-10 lost total hydraulic power as a result of a mid-flight tail engine rupture, with debris ripping through all hydraulic lines that ran through the tailplane. The crew figured out how to use differential power on the two remaining engines and steered the craft toward an extremely
difficult high-speed landing at Sioux City, Iowa. In spite of that the plane broke up on the runway a majority of the passengers and crew subsequently survived the landing.

Thinking outside the box, taking a system way beyond what it was designed to do (even making use of an adverse design quality such as pitching moments with power changes), are hallmarks of resilience. Resilience is the ability to recognize, absorb and adapt to disruptions that fall outside a system’s design base (Hollnagel, Woods, & Leveson 2006), where the design base incorporates soft and hard aspects that went into putting the system together (e.g., equipment, people, training, procedures). Resilience is about enhancing people’s adaptive capacity so that they can recognize and counter unanticipated threats. Adaptive capacity with respect to a narrow set of challenges can grow when an organization courts exposure to smaller dangers (Rochlin, LaPorte, & Roberts 1987). This allows it to keep learning about the changing nature of the risk it faces—ultimately forestalling larger dangers. Such adaptation could be one explanation behind recent data that suggest that the passenger mortality risk on major airlines that suffered non-fatal accidents is lower than on airliners that had been accident-free (Barnett & Wang 2000).

United Airlines 232 at Sioux City is such a case. Crews can counter many threats effectively by replicating or slightly varying the technical skills learned during their training. Most non-normal situations in commercial aviation, after all, are quite ordinary or at least recognizable: they fit within the “box” and, accordingly, crew behaviour can stay “inside the box.” Then there is a huge middle ground. It consists of daily safety threats that feature (or occur because of) subtle variations that call for some adaptive capacity and response. For example, these threats can demand extra work (e.g., gathering and processing of more data, increased communication and coordination), recruitment of additional expertise (e.g., dispatch, ATC) and the deployment of new strategies. Resilience, however, means effectively meeting threats that represent infinite reconfigurations of—or ones that may lie entirely beyond—what the industry could anticipate.

2. Simulation fidelity and development of resilience

The aviation industry has relied on simulation perhaps more than any other safety-critical industry. While simulators are still used for stick-and-rudder and instrument training, they are also today part of practically all aspects of aviation training (Salas, Bowers & Rhodenizer 1998) and the new Multi-crew Pilot License (MPL) rests almost entirely on simulated flight training. This investment in simulation reflects an industry-wide confidence that it can save time, money and lives (Bürki-Cohen, Soja & Longridge 1998), not to mention provide effective training—i.e. develop skills and knowledge which are transferable to any target situation.

The evolution of simulation in aviation has mainly been technology-driven, from the introduction of visual systems and computer graphics (Dennis & Harris 1998, Lee 2005), to recent additions of satellite imagery to represent the visual scene of the ground below, and current moves to further integrate into simulations the role of air traffic control (Longridge, Bürki-Cohen, Go & Kendra 2001). Although this increase of face-validity to improve training quality has been questioned (Roscoe 1991, Salas, Bowers & Rhodenizer 1998, Dahlström & Nählinger 2007) there seems to be a taken-for-granted assumption within the aviation community that incremental quantitative progress (e.g. more computing power, higher resolution, greater visual angles) adds up to a positive qualitative difference. In other words,—as a simulated environment becomes ever more “photorealistic”, so does the yield simulations have for crews and staff. This link between maximum fidelity to maximum training transfer is taken on faith: if it looks real it will provide good training. However, the continually increased demand for higher levels of fidelity to make simulation look real leads over time increases cost and lowers availability of training simulators (one has to keep in
mind the huge capital investment a commitment to this style of simulation requires). Even though there have been few studies of transfer of training from photorealistic simulators to aircraft (Carretta & Dunlap 1998, Dahlström & Nählinder 2007) and the problems of performing such studies have been documented (Bell & Waag 1998, Hays, Jacobs, Prince & Salas 1992) the assumed relation between fidelity and transfer of training seems to prevail in the aviation industry (in other industries too, e.g. maritime transport, nuclear power, medicine, as well as in the military).

Some have suggested that lower-fidelity simulations (ones that do not attempt to mimic directly the target technical environment) are a cost-effective alternative and may actually improve many aspects of learning that help people deal with unanticipated situations (Rouse 1981, Roscoe 1991, Caird 1996). As Caird (1996, p. 127) put it, “For decades, the naïve but persistent theory of fidelity has guided the fit of simulation systems to training.” In addition, Heeter (1992) concluded that the environmental presence experienced in simulated environments is determined more by the extent to which it acknowledges and reacts to the participant than by the simulation’s physical fidelity. In other words, high levels of technologically-driven fidelity can simply be wasteful in terms of costs and time relative to the pedagogical undertaking at hand. (The cost of these simulations may also inadvertently limit access to what the industry considers to be “proper,” “legitimate” training opportunities.) As well, “featurism” can be distracting (Jackson 1993), both for the trainer and the trainee, especially when the features argued for, promoted and designed in are skewed in the direction of realism.

The emphasis on photorealism in visual and task contexts may retard or limit the development of skill sets critical for creating safety in domains where not all combinations of technical and operational failure can be foreseen or formalized (and for which failure strategies then cannot be procedurally modified and simulated). The assumption that photorealism can capture all possible naturalistic cues in addition to the skills necessary to act competently in these domains may be overly optimistic. Competencies the aviation community recognizes as important and significant (e.g. communication, coordination, problem-solving, management of unanticipated and escalating situations) are thought to emerge directly from context-fixed simulator training. It is assumed that photorealism can achieve these ends.

The focus on face-validity has muted perspectives on simulation styles and use that could allow a more subtle analysis of cognitive and group interaction aspects to form the base of training. This is particularly true for training of skills related to Crew Resource Management (Baker, Prince, Shrestha, Oser & Salas 1993). It is in unusual, unanticipated and escalating situations where such skills are most needed. Highly dynamic situations involving underspecified problems, time pressure constraints, and complex group interaction are situations that cannot be resolved through procedural guidance. Pertraglia (1998) observed that experience in the world can be neither “predetermined nor preordained” and that this, together with the willing suspension of disbelief, is what should make a simulated activity seem authentic as experience. However, the quarter-century long aviation industry focus on Crew Resource Management has resulted in few attempts to provide training that addresses situations with underspecified problems and time pressure in the context of group interaction. The commitment to highly realistic simulation has meant that crews have not been well trained when it comes to situations which are neither “predetermined nor preordained”. The recently introduced Multi-crew Pilot License (MPL), which aims to qualify candidate airline pilots as part of a crew from the very beginning of their training by increased use of simulated flight training, is an opportunity to review the relationship between simulator fidelity, quality and transferability of training and the underlying assumptions this training is based upon.
3. Fidelity and validity: a more complex relationship than assumed

At Lund University School of Aviation, we have been experimenting with lower-fidelity simulations over the last two years, to assess the relationship between simulation fidelity and validity—that is, the connection between the faithfulness of the constructed world in which operators are trained on the one hand, and the extent to which this actually supported the development of skills useful in target environment.

We have used a simulation of a ship’s bridge (essentially consisting of a laptop computer, printer and table top) that is simple in regards to the participant interface (printouts) but with time-pressure and various event-driven scenarios built-in. These range from minor problems to serious threats to the safety of the ship M/S Antwerpen (Strohschneider & Gerdes 2004). This mid-fidelity (far from “real” but capturing sufficient and salient aspects of reality) simulation has been run twice during a two-day M/S Antwerpen training program, which also included debriefings, discussions and lectures. The participating groups were made up of student pilots, maritime students and operators and the safety and security group of Lund University. Observational notes and video recordings were taken to allow the simulation sessions to be used as a series of case studies on information management, group coordination, leadership and decision making as well as on management of unusual, unanticipated and escalating situations (Dahlström, 2006).

During debriefings participants have consistently stated that this simulation provides them with relevant and valuable training for their actual work and explicitly wished for more training of this type. In particular, participants with no maritime background find the “cross-domain” elements of the training beneficial. In spite, or perhaps because, of its lack of fidelity to photorealistic representation and feedback, the engagement and level of intensity of communication, cooperation and decision making observed in groups normally surprise the participants themselves as well as instructors. Participants typically bring this up and note that it has had as an important influence on the training effect. There have been very few requests from participants for increased fidelity to improve the simulation; in fact contrary opinions have been more often recorded. These comments, together with observations of the over-emphasis (in particular on the first simulation run ) on technical “real world” ship parameters (e.g. engine RPM, course and roll angle) and lack of process-oriented discussion, indicates that potential “improvements” to fidelity could in fact have a detrimental effect on the validity of the training. For example, if we were to “engineer in” higher-fidelity features, e.g. knobs, levers and buttons participants are of the opinion that this would shift their focus from generic to procedural competencies. Also, since building in such higher-fidelity items could never bridge the qualitative gap to the “real” thing in any case, much participant attention and commentary would be directed to their insufficiency or still unconvincing nature. The lack of such features leaves groups with no option but to focus on use of general competencies as tools to manage the situations they encounter.

Participant concerns did not focus on the relevance of the simulation (most participants, it needs to be kept in mind, were non-maritime operators) or on issues during simulation exercises related to maritime technical or procedural knowledge. Instead we heard much from participants how training of this kind could support development of competencies and understanding applicable to problems operator in any industry might encounter. To experience and to learn how to address problems such as unusual, unanticipated and escalating situations in combination with underspecified problems, time pressure constraints and complex group interaction (key features of the simulation) participants recognized as important but these issues were ones they seldom encountered in regular operator training. Our observations support that a shift of domain seemed to recreate and emphasise the types of uncertainties reported in incident and accident investigations. High-fidelity flight simulator training is normally focused on removing, rather than enabling participant understanding, of such uncertainties. In addition, the non-domain specific environment seemed to encourage
participants to step out of their normal roles and explore aspects of general group interaction competencies – ones not covered by standard procedures and instructional theory. An example of this was seen during a fire on-board the M/S Antwerpen, where one operator in a group of maritime operators insisted on relieving the captain of handling passenger evacuation since he was getting overwhelmed trying to control the spread of the fire. Also, an analysis of M/S Antwerpen group processes and outcomes (in terms of saved lives and damage to ship) indicate that groups which built up and relied on generic competences performed better than those who relied more heavily on established roles and procedures (van Winsen, in preparation). This is most clearly seen when maritime students, on M/S Antwerpen exercises cooperated far beyond the formal responsibilities of their respective roles compared to maritime operators, who established standard roles and explicit procedures that they then were unable to break out of as an emergency escalated.

It may be that problem-solving exercises carried out in high-fidelity (realistic), highly context-specific environments, can impede people’s imaginative, creative involvement and the resilience it may eventually delivers to the workplace. Training in high-fidelity settings alone valorises the internalisation of a series of highly contextualized instrumental stimulus-response relationships—putatively stress-resistant procedural responses which may be insensitive to, or even make actors unprepared for, contingencies outside of rehearsed routines. If the desire is to have operators successfully extrapolate and improvise beyond a set of fixed learned responses, this issue of what is “carried away” from context-specific simulation exercises needs to be looked at more carefully than it has in the past.

As Roscoe (1991, p. 1) notes, “Research has shown that innovations in training strategies, in some cases involving intentional departures from reality, can have stronger effects than high simulator fidelity on the resulting quality of pilot performance.” Indeed, as Caird adds (1996, p. 128): “…there is some evidence from flight simulation that higher levels of fidelity have little or no effect on skill transfer and reductions in fidelity actually improve training. Reductions of complexity may aid working memory and attention as skills and knowledge are initially acquired.” It may be that the lack of physical fidelity in the lower-fidelity simulation can enhance the focus on training of general principles of communication, coordination and problem solving in a workgroup. These principles are ones actors can use to understand and resolve situations beyond those covered by procedural guidance (Dörner 1996).

Locking training to context-specific environments affects more than the exportability of instrumentally-rehearsed skills. It can also amplify and reify role socialisation. Effective management of escalating or otherwise novel situations has been associated with the breaking-out of roles and power structures that were formally designed into the system. It is not clear whether naïve (built to mimic reality) simulation can “train” in this direction at all. When roles are involved Weitz and Adler (1973, p. 224) concluded that “…it might be wise to stress the principles, not the roles” to ensure that participants do not “become wedded to particular performances”. Roles and power structures often go hand-in-glove (e.g., Captain and First Officer), and various programs (e.g., Crew Resource Management training in aviation) aim to soften role boundaries and flatten hierarchies in order to increase opportunities for coordinating viewpoints and sharing information. Operational success in the face of extreme or rapidly shifting demands can hinge on individuals going beyond the formal roles assigned to them—as illustrated by various near-accidents or accidents that could have been worse, such as United 232 at Sioux City (and see also Dekker 2005).

4. Conclusion

If no training opportunities exist in which individuals can disconnect from the constant reification and elaboration of their normal operational activities, the ability to respond effectively may remain inextricably anchored to (and fundamentally limited by) known and
rehearsed roles, duties and procedures (as shown powerfully in Weick 1993). Lower-fidelity simulation can serve as an important resource in the creation of resilient operators. In the research we have conducted, it seems to force participants to confront the interpersonal and goal-oriented demands of managing unanticipated and escalating situations. These are exactly the work elements that seem to be lost or hidden behind the procedural specifics fostered by high-fidelity training and operational experience. In addition, the limited costs connected to using lower-fidelity simulation could increase availability and frequency of training sessions.

Our case study of accident cases and use of lower-fidelity simulation seems to reveal a disconnect between fidelity (or photorealistic faithfulness) of a simulation and its validity (how the skills it develops map onto situations in the target environment). Lower-fidelity simulation allows the development of generic problem-solving skills such as sharing knowledge, making and following up on plans, dividing work, stepping back for broader evaluation, borrowing time from the future by current task investments, and maximally exploiting a group’s available expertise. These skills (and the confidence that comes from successfully deploying them in settings other than the target environment) could contribute significantly to the development of resilient crews in ways that considerably more costly and more high-fidelity (photorealism) training cannot. Traditional assumptions about simulation tend to portray both role and context as though they are natural, unalterable facts. This message seems to be implicit in almost all attempts at “realistic” simulation. We would, however, argue that this conveys exactly the wrong message if we want individuals and workgroups to be adaptive and capable of creative, appropriate improvisation—skills which are practiced and learnt effectively in lower-fidelity simulations. These are simulations which by design can lead participants to rethink their normal roles and behaviour. This in turn can help develop more adaptive and flexible competencies which can strengthen operator and system resilience in the face of unanticipated and escalating situations.

Acknowledgments

We gratefully acknowledge our cooperation with Professor Stefan Strohschneider at the University of Jena in Germany who, together with Jürgen Gerdes, originally developed the M/S Antwerpen mid-fidelity simulation and training program. The research presented in this article was funded by the Swedish Emergency Management Agency and the Swedish Governmental Agency for Innovation Systems.

References


Learning from failures in emergency response: A series of empirical studies

Research paper

Sidney Dekker*, Magnus Jonsén, Johan Bergström and Nicklas Dalhström

Lund University
TFHS
Drottningvägen 5
SE-260 70 Ljungbyhed
Sweden
Phone: +46-435-445434
Fax: +46-435-445464
sidney.dekker@tfhs.lu.se

*corresponding author

Learning from failures in emergency response: A series of empirical studies

Abstract

Purpose: Recent high-visibility disasters have fuelled public and political awareness of the importance of managing and mitigating their consequences effectively. In response, various countries have enacted legislation that demands the evaluation of emergency responses so that lessons for improvement can be learned. Methodology: We conducted a series of field and experimental studies from 2005 to 2007 to assess the ability of first-responder organizations (e.g. fire departments) to learn from failures that occurred during their emergency responses. Findings and value: The departments we studied often lack basic organizational requisites for effectively learning from failure (e.g. mutual trust; participation, knowledge of possible learning mechanisms). Further, neither first-responder training, nor daily practice, seems supported by knowledge of generic competencies necessary for effective crisis management. This not only hampers coordination during a response, but also keeps its evaluation from using a language that could help organizations learn and improve.

Keywords: learning, failure, emergency response, safety culture
Introduction

Recent high-visibility disasters have fuelled public and political awareness of the importance of managing and mitigating consequences of disaster effectively (see e.g. OECD, 2003). In response, various countries have enacted legislation that demands the evaluation of emergency responses so that lessons for improvement can be learned. First-responder organizations have come up with a variety of ways to comply, ranging from self-evaluation sheets, checklists, documentation of de-briefing sessions, to formal inquiries of the emergency response by a dedicated evaluation person or team. The requirement to assess emergency responses often comes on top of other stipulations, such as having to investigate the causes and sequence of events of the incident that triggered the response. As a result, the resources for, and the quality and extent of, evaluating emergency responses varies widely. Checklists and self-evaluation sheets can quickly join a compliance paper-trail that leaves no meaningful footprint on the organization or its core mandate. The enthusiasm with which emergency response reports get written is seldom matched by the willingness, or time taken, to read those reports, let alone implement any of their recommendations. Indeed, first-responders often have few effective ways of disseminating lessons learned beyond their own organization or even beyond the person or team involved (Tranqvist, 2004; Sellberg & Dekker, 2006).

It would seem that the ability to learn from failure in emergency response can be improved in obvious ways. One would be to enhance the technical mechanisms for collection and dissemination of data about a response (e.g. computer-based collection and analysis, web-based publication). Other ways would be to assist team leaders during debriefings or to standardize the way in which analyses are conducted and written up (such as is done with aviation accidents according to rules from the United Nations’ International Civil Aviation Organization). Yet others would involve more formal requirements for the follow-up and implementation of recommendations that come out of an evaluation. In the course of our work, however, we discovered that such measures hardly have a chance of success if there is no serious consideration of more fundamental barriers to learning from emergency responses. Some of these barriers seem related to cultural predispositions and organizational relationships, while others could be the result of the focus of training and proficiency-checking in first-responder work.

Method

In a series of field studies, we followed the emergency responses and evaluations of one rural and two major urban fire departments in Sweden during a period of six months. Through a combination of participant observation and targeted interviews, we were able to assemble an inside-out view of what it means to learn from failure as first responder, and what the organizational constraints and opportunities for learning are. All departments studied were (in part because of regulatory requirements) working on implementing some kind of learning process, ranging from self-rating scales to appointing an investigator. Field study results motivated us to more formally test the ability of first responders to reflect on their group’s performance after its management of an escalating situation. We were particularly interested in experimentally assessing the generic competencies that first-responder leaders take with them into action, and that they will rely on in the evaluation of their team’s performance (including its failures). We divided a class of final-year (of a 4,5-year program) fire commander students in two. Half the class was offered a two-day intensive training program for managing escalating situations, which involves a simulated passenger ship bridge where multiple threats have to be managed over a number of hours (rough seas, water penetration, onboard fire; see Dahlström et al., 2007). Trainees were given roles typical
of the bridge crew (Captain, Chief Officer, Chief Steward, Doctor, two Engineers). As part of the training program, their ability to handle escalating situations was evaluated and discussed extensively. We then conducted a follow up during the class’s own field exercises, in which senior students played the roles of their normal domain. Using the results of these exercises, we were able to compare the group of students who had received the training in generic competencies of emergency management with the other half of the class.

**Results and discussion**

Whether in the field, on the way back in the fire truck, or across the management conference table, people and groups reflect on their emergency response work. And they occasionally make changes based on their idea of how to improve or fine-tune aspects of that work. This does not necessarily mean, however, that these reflections and changes are exhaustive, effective or even constructive. In fact, a number of factors seem to conspire against first-responders’ ability to extract and disseminate meaningful lessons from their emergency work. *Human error as target for improvement.* One typical response to failure during an emergency response is to blame “human error” by any other name (e.g. complacency, did not watch out, judgment failure, inadequate supervision, incompetent leadership). This folk model, where failure is caused by human errors, is often firmly in place, leading stakeholders to suggest that the reprimanding or compensatory training of individuals is a plausible countermeasure. Alternatively, first responders can be reminded of applicable rules, routines, procedures or regulations and asked to more stringently adhere to them. This is not necessarily a fruitful way of engaging highly variable contexts (see Dekker, 2003). A final countermeasure based on this folk model is to supply more technology (e.g. communications- or information technology). While offering new capabilities, these solutions often have unanticipated side-effects. They create new work for commanders and their crew (e.g. managing the interface, remembering codes), and can lead to communication clutter and data overload for those in command roles (Woods, Patterson & Roth, 2002). New technology rarely solves organizational or command problems, but rather articulates or changes their expression (Suparamaniam & Dekker, 2003).

Yet the need to redirect improvements away from “human error” seem counterintuitive if the problems caused seem so indissolubly linked to human performance weaknesses. One way forward for first-responder organizations is to see human errors as consequences, as effects or symptoms of deeper trouble, rather than as causes of trouble (since the latter quickly blocks learning). What is interesting is not the number or category of errors, but the complexity of the operational work in which they occur. This can allow organizations to start seeing commonalities and patterns across responders, teams and situations. Our field work consistently confirmed how people in these jobs try to create safety under uncertainty, time pressure, resource limitations and multiple performance goals. Failure is the inevitable occasional by-product of the pursuit of success under those circumstances, and those failures can often be systematically traced back to features of people’s operational, technical and organizational work environment.

Extracting learning from those failures, however, does require a particular kind of vocabulary and tonality for discussing emergency response failures with those involved. We noted entire repertoires of defensive posturing, when commanders or fire-fighters felt that they or their assessments and actions are under scrutiny or attack. These included shooting the messenger, valorizing operational experience (“we know how to put out fires, who are they to question our judgment”), and even appealing to masculinity or its supposed dearth. We did see, though, that meaningful contemplation is possible without invoking defense mechanisms. For
example, during a debriefing, a commander began humbly by addressing his own misjudgments and errors before asking the entire response team to reflect. Such openness, however, is often limited to people’s own organizational groups or levels, and does not extend to outside probing.

**Trust and participation.** Critical for the success of any learning that involves the sharing of stories of failure is the existence of trust between the parties who disclose and who receive the story. In other operational worlds (e.g. aviation), such trust is engineered into the system in various ways. One is anonymity (not having to report a name or any other identification, only the substance of the story) or confidentiality (having any identifiers or names taken out of the story before it is sent on). This may not be feasible inside individual emergency organizations, since their often limited size will enable people to identify one another regardless of such precautions. But it could work on a regional or national scale. In fact, in those cases, having neutral outsiders gather, assess and synthesize could contribute significantly to people’s willingness to share stories of failure. Reporting outside the line organization (either into a local staff, or to a neutral outsider) is particularly helpful if people have concerns about those they work with (especially their superiors) or if they fear consequences for their own careers or work situations as a result of their reporting (Steele & Dekker, 2004).

Trust may remove barriers for people to share and report stories. But trust is not enough to actually get people to share or report. Participation in change efforts is critical to motivate people to report and keep reporting (Kohn, 1999; Dekker & Laursen, 2007). We saw how one large fire station took away the idea-letterbox that they had earlier created to trawl their operational personnel for improvement proposals. Despite generating a number of nifty solutions to local problems (e.g. mounting a first-aid kit in the top of the ladder truck), the system was taken away because it did not match bureaucratic accountability requirements for quality improvement work inside the organization. This had the halo-effect of relaying to firemen that their contributions to learning were not sufficiently worthwhile. It measurably dented their motivation to participate in subsequent improvement or change. Similar actions, that were perceived as put-downs, included using the word “whining” to describe how local personnel considered certain improvements in their working conditions necessary.

In contradistinction, and consistent with the literature (e.g. Kohn, 1999), organizations that were able to harness the knowledge and experience of those on the front-line of their safety-critical work in improvement efforts often had little trouble filling their reporting systems with meaningful stories and suggestions (Dekker & Laursen, 2007). Ways in which they did this included having peers debrief operators and review with them their reports and improvement proposals, and sending feedback into the organization about what was accomplished as a result of those suggestions. That said, getting reports is not the same as learning anything of value from them. Documentation can never double as learning. This is where effective reporting systems can succumb to their own success: their ability to invite reports oversteps the analytic ingenuity and synthetic abilities to see larger patterns that offer leverage for change. The only organizational answer there is to increase the resources spent on analyzing and synthesizing reported information.

**Not offering a second opinion.** In many cases, the evaluation of emergency responses amounted to people grading their own homework. This can be quite useful, but we saw limits where there were fears that negative self-grading would hamper one’s organizational stature or career prospects. On many self-assessment sheets individuals filled in “no comment,” rather than reflecting critically on (and negatively grading) their own performance. “No comment” of course means no learning. A second opinion could be introduced by having others from outside the team or even organization look at and assess the evaluations and ask...
the kinds of questions that can only be asked if one is not part of the culture that takes its particular practices for granted, as normal, unquestionable. This does not have to be more formalized than rotating individuals through different departments or across organizations so that they can bring a fresh look or independent perspective onto something that insiders may no longer see as either problematic or open for change. Independent safety investigations are of course deeply institutionalized in other operational worlds (e.g. aviation), which generates a whole different depth of failure analysis (though not automatically an easier acceptance and implementation of any recommendations generated).

Fine-tuning the margins of existing practice. Closely related to the issue of second opinions, is that much of the observed learning seems to take place at the margins of existing practices. Routines are amended slightly (e.g. decisions on placement of a ladder truck), but the substance of existing practices is hardly ever challenged (e.g. sending a ladder truck to a traffic accident, dispatching a commander plus four fire-fighters to every call, prioritizing the sending in of fire-fighters over attacking a fire from outside) even if they are insensitive to subtle variations in context, or based on obsolete assumptions about operational work or long-since disappeared technologies. Critically questioning the very basis for some of the operational work that groups engage in when responding to disaster seems much more difficult than making some fringe adjustments, and its possibility or necessity does not even occur to most of the people we worked with or interviewed. This is, once again, where routine independent investigations of emergency responses could pay dividends. Their independence, however, (and thereby ability to look at established practice from a fresh viewpoint) needs to be balanced by a credibility in the eyes of the recipients of any conclusions and recommendations. That can be done only if the independent investigation integrates a sufficient level of technical expertise (which in turn can challenge its independence; its ability to withstand inculcation and then taking particular practices or factors for granted).

No language for generic emergency management competencies. Perhaps the most fundamental barrier to learning was the lack of an awareness of, and even language for, generic competencies for emergency management. This lack first became apparent in debriefings or other performance evaluations that we observed. Such general competencies are generally assumed to include information management (sorting, prioritizing, explicit goal statements), leadership and communication (flexibility, building up shared mental models), proactive strategies and analysis, and assessment of any intervention effects and revision of plans (e.g. Dörner, 1996). None of these notions were explicitly taken up in how groups talked about their performance during an emergency response, even though failures that occurred could frequently be reduced to precisely such cognitive and coordinative problems.

This motivated us to test whether there is evidence of an awareness of such competencies in senior student-commanders. During the first simulated fire, trainee-commanders behaved like groups whose professional indoctrination or life does not revolve around emergency response. They displayed a quick glide into “normal operations” (not preparing for anything unusual or unexpected) even though it was clearly stated that they were to perform a crisis-management exercise with a simulation method and that they should make no assumptions about the condition of the ship. There were no efforts at pro-activity, no explicit goal formulations, an inability to sort through escalating information flows, few attempts at process discussions with the emergency management team, no strategies for distributed decision making, and no flexibility in role execution.

Even though the students’ generic competences for crisis and fire management was indistinguishable from those who did not spend the past four years being educated as fire commanders, they responded well to the two-day training program. The group that underwent
the training program in handling escalating crisis performed much better in their subsequent field exercises than the other half of the course. Roles were better defined, including those of team leader and moderator. Not all decisions were taken in consensus (nor was that judged necessary), decisions got clearly executed and followed up, and there was a modicum of explicit goals that the team strived for. Interestingly, discourse about their performance both during and after the field exercises, became more articulated in the groups who had gone through the training program in generic competencies, and focused on the social and interactive aspects of team performance—those aspects that alone can carry success or failure in managing an escalating situation. For all trainee-commanders, this was the first time that the social and interactive aspects of performance were taken up explicitly at all, their training having otherwise focused entirely on technical competencies and operative competencies on a rule-based level.

**Conclusions and recommendations**

Learning from failures that occur during an emergency response cannot be improved by narrow organizational or technical solutions alone, if at all. Better mechanisms for reporting and disseminating information can work successfully only if founded on organizational trust and participation in improvement processes. Also, learning from failures during an emergency response requires a particular language that not only forestalls psychological defensiveness, but is capable of bringing out the social and interactive aspects of team performance that are critical to the success or failure in managing an escalating situation. Based on the empirical work described in this paper, we conclude and recommend the following:

- There seems to be a need to evaluate emergency response training curricula, particularly those for commander training, to see to what extent they integrate the teaching of human factors issues that are of great importance for the success of emergency responses (e.g. communication, coordination, situation awareness, delegation, information management, creation of a safety culture). In addition, such curricula could include an early focus on what organizational learning is, and how to promote it.
- The development of general competencies for emergency management such as information management, group interaction, leadership and decision making could be supported effectively by putting trainee-commanders in contexts other than those where they can rely on rehearsed roles and over-learned routines. Evidence from the work reported here suggests that occasionally setting practitioners outside their normal context, may accelerate the development of awareness and discourse around generic skills for handling escalating situations.
- An earlier integration of the training programs of those who will be working at different levels in the emergency response organization’s hierarchy (e.g. commanders versus firefighters) could promote the kind of understanding, trust and openness that is necessary for organizational learning later on.
- Documenting an emergency response by whatever means (formal report, checklists, self-assessment sheets) is not the equivalent of learning from that response. There is a strong need for resource investment in organizations’ analytic and synthetic capabilities so that such documentation can be transformed into lessons-to-be learned, and disseminated as such.
- Trust is important if organizations wish to remove communication barriers so that people start sharing and reporting stories of failure. Actually getting people to report requires enhancing people’s participation in the learning process. There should be
interest in and respect for ideas that come up from operational personnel on how to improve the quality of the processes they work with everyday.

- The institutionalization of a second opinion during learning from emergency response failures, varying from rotating one’s own personnel to delegating investigations of emergency responses to an external, independent organization, can be a critical ingredient in questioning and improving practices and routines that could otherwise remain taken for granted.

References


Appendix D. Preliminary version of article

The Limiting Effect of Procedures on Resilience

Roel van Winsen, Nicklas Dahlström, Sidney Dekker\textsuperscript{2}, and James M. Nyce\textsuperscript{1,2}

\textsuperscript{1}Lund University School of Aviation, Drottningvagen 5, 260 70 Ljungbyhed, Sweden

\textsuperscript{2}Department of Anthropology, Ball State University, 2000 W. University Ave., Muncie, IN 47306, USA

1. Introduction

1.1 Procedures & Resilience

In today’s world, the aviation industry is widely regarded as a very safe industry. In its short history it has gained this reputation due to the continuing evolution and success of systematic implementation of standardized procedures for all aspects of operations. Use of procedures, supported by checklists and manuals, to carry out operational activities increases reliability of human behavior and makes it predictable to designers, regulators, operational management and instructors.

Our current belief in standardizing behavior through procedures can best be seen when looking at reactions to failure of the system: more procedures are introduced or more strictly enforced in order to control the weakest chain in the shackle; the unreliable non-compliant human operator. This is illustrated by the reaction given shortly after a fatal shootdown of two US Black Hawk helicopters over Northern Iraq by US fighter jets: “higher headquarters in Europe dispatched a sweeping set of rules in documents several inches thick to ‘absolutely guarantee’ that whatever caused this tragedy would never happen again” (Snook, 2000, p. 201).

Operational activities in any industry, especially in the transportation sector, often involve situations whose subtle and infinite variations mismatch the predefined procedures. Emergencies are such situations that most often do not result of foreseeable events but occur due to unpredictable interactions in complex systems (Perrow, 1984). In the complex socio-technical environment of an aircraft, for example, there will always be a residue of potential problems that crews cannot be prepared for, i.e., they are not in their inventory (see Dismukes, Berman & Loukopoulos 2007). Formal mechanisms of safety regulation (through e.g., design requirements, policies, procedures, training programs, line checks) will always fall short in foreseeing and meeting the shifting demands posed by; uncertainty, limited resources and multiple conflicting goals.

As a result, some crews will at some point be left to “fend for themselves” at the edges of a safe industry, as operators must be able to apply skills and knowledge that no training or standards department was able to foresee. For this residue we have to rely on crews’ generic competencies in order to be resilient. That is, when operators find themselves in a situation for which no procedures apply, we first want them to recognize this, then analyze the problem, be creative and solve the situation accordingly. A true example of resiliency can be seen in the flight of United Airlines 232 (National Transportation Safety Board 1990). The DC-10 lost

\textsuperscript{2}Corresponding author. Email: sidney.dekker@tfhs.lu.se
total hydraulic power as a result of a mid-flight tail engine rupture, with debris ripping through all hydraulic lines that ran through the nearby tail plane. The crew quickly adapted and learnt how to use differential power on the two remaining engines and steered the craft toward an attempted landing at Sioux City, Iowa, which a large number of passengers and the crew subsequently survived.

Thinking ‘outside the box’ and thereby taking a system far beyond that what it was designed to do (even making use of an adverse design quality such as pitching moments with power changes) are hallmarks of resilience. Resilience is the ability to recognize, absorb and adapt to disruptions that fall outside a system’s design base (Hollnagel, Woods, & Leveson 2006), where the design base incorporates soft and hard aspects that went into putting the system together (e.g., equipment, people, training, procedures). Resilience concerns the enhancement of people’s adaptive capacity so that they can recognize and counter unanticipated threats.

United Airlines 232 at Sioux City is an example of an extreme case. In general crews can counter many threats effectively by replicating or slightly varying the technical skills learned during their training. Most situations in commercial aviation, after all, are quite ordinary or at least recognizable: they fit within the ‘box’ and, accordingly, crew behaviour can stay ‘inside the box’. Then there is a large middle ground consisting of daily safety threats that feature (or occur because of) subtle variations that call for some adaptive capacity. For example, these threats can demand extra work (e.g., gathering and processing of more data, increased communication and coordination), recruitment of additional expertise (e.g., dispatch, ATC) and the deployment of new strategies. Resilience here means effectively meeting threats that represent infinite reconfigurations of—or that lie entirely beyond—what the industry could anticipate.

The question now becomes whether we can expect this kind of resilience (stepping back to look at the problem from a different, more open, viewpoint) from operators, experts as it may be, who are trained to follow the rules, adhere to checklists, and who are told/believe that there is a procedure for every occasion? As an example of how difficult it is for crews to recognize that procedures do not apply any longer, we can examine the Swissair 111 accident (Transportation Safety Board of Canada 2003). During this accident the presence of smoke in the cockpit was responded to promptly by the crew by adhering to the relevant checklist. However, while following established procedures for this situation, the fire engulfed the aircraft. In this case following the procedures turned out to be the problem rather than the solution (see Burian & Barshi 2003).

There are hardly any training programs that train operators how to be, or better behave, resilient. The idea seems to be that by training operators for ‘normal’ kinds of emergencies and giving them all sorts of procedures to adhere to under these circumstances, operators will also be better able to deal with all emergencies (including those unimaginable and unprepared for) thus becoming more resilient automatically. In this view, procedures and such specific training are assumed to have a causal effect on resilience. In this study we set out to explore this relationship by looking at the effect of the availability of procedures on operators’ ability to be resilient.

1.2 The M/S Antwerpen simulation

The M/S Antwerpen simulation is part of a two-day emergency management training course developed by the University of Bamberg (Strohschneider & Gerdes, 2004). However, apart
from its training purposes, the simulation can be used as a research tool for providing data on group action and interaction in escalating events.

The simulation is designed for a group of five to seven participants who act as the ship’s first officers. Each participant takes on a specified role, namely: the captain, the first officer, the first engineer, the chief steward, the ship’s doctor, the main engineer, or the navigation officer. Initially each participant is provided with general information which describes the features of the ship as well role specific information. It is the participants’ task to safely navigate the ship through a stormy night in the North Atlantic. Due to the adverse conditions, and because the ship has been poorly maintained, the crew is forced to deal with a number of passenger-related problems and technical failures that towards the end of the simulation result in a state of emergency.

To sail the ship and handle these events, the participants have a wide variety of options and actions available to them. They have control over the technical facilities of the ship, including maintenance and repairs. Furthermore, they are presented with an abundance of information regarding the ship. They can direct the crew and give various orders relating to the passengers (including, e.g. sending misbehaving passengers to their cabins, closing or evacuating sections of the ship, and using life boats and rafts to abandon the ship). Participants are not provided with a prescribed list of possible responses. Instead they have to plan and execute actions as a team and deal with the possible consequences and side-effects of their actions. The participants, therefore, find themselves in a dynamically developing situation, one that has a high level of uncertainty. Moreover, they have to deal with all this under the threat of all the conceivable emergencies that come with navigating a poorly maintained ship in bad weather.

1.3 The effect of different levels of procedural-experience

In this study, three different types of groups participated in the M/S Antwerpen training program; each group differed in their familiarity with maritime operations, and therefore their availability of applicable procedures. We were interested in the effect of these different levels of experience and available procedures, on various concepts regarding team interaction and management of escalating situations.

The first type of group existed of novices in relation to maritime settings, namely civil aviation student pilots. This group had minimal knowledge of maritime concepts, operations, and procedures. The second type of group consisted of maritime students with limited experience of maritime operations. The members of this group were familiar with most maritime concepts, however they did not have any practical first hand experience with the practices and procedures on large ships. The third group consisted of experienced seafarers, who had multiple years of experience on large ships. They were therefore very familiar with maritime concepts, as with the normal and emergency procedures and practices onboard such ships.

Our main point of interest in this study was whether or not the availability of procedures diminished the crews’ generic competencies and therefore their abilities to be resilient. In order to approach this broad issue we observed various general and specific aspects of the participants’ individual and group behaviours. When observing the three groups in order to assess their generic skills, we were interested in the following concepts; team cooperation and communication, management of workload, and information sharing.
Firstly team cooperation and decision making was specifically examined as we recorded whether or not the participants: openly stated their personal and group’s goals, whether or not these goals were discussed, and whether or not decisions were based upon them. Furthermore, we analyzed the manner in which decisions were made, looking specifically at: whether or not initiatives were first discussed within the group or within subgroups, or alternatively blindly initiated by the individual team members.

Secondly we focused our attention on how team members dealt with workload and role definition. We were interested to see if participants shared their tasks and workload when this was necessary, or whether they rigidly remained within the work constraints of their prearranged roles. Furthermore, we were interested in whether people would step out of their roles and challenge or take over other team members’ responsibilities and make decisions accordingly. How the groups dealt with the idle time during the simulation was a further area of interest. Idle time refers to the periods during the simulation in which there is hardly any work to do for the group. Idle time provided the groups with the opportunity to think ahead, proactively try to prevent problems, and construct team strategies accordingly, instead of just waiting for the next crisis to come along.

Thirdly, in regards to information sharing we looked at the manner in which the groups solved the information overflow problem. The simulation ensures that the groups are required to find a way to manage the continuous printer outputs. How this was approached and thereby how incoming information was broadcast to the group or its individual members was examined. We observed and recorded the manner in which participants shared information by means of briefing routines or how they designed other structured approaches. Finally, but perhaps most important, we examined the groups’ overall capabilities to step back and on a meta-level analyze, discuss, and adjust their approach and team functioning during the simulation.

It seems that the different levels of (procedural) experience described above give rise to different ways of managing emergencies and escalating situations in regards to the above mentioned concepts. In analyzing the performance of the groups, we distinguished between the perspectives of ‘process’ and ‘outcome’. Outcome relates to the quantitative results of the simulation; primarily numbers of injuries and casualties, and damage to the ship. As a performance measure, however, the outcome of the simulation is dependant on the interaction between the participants and the facilitators of the simulation and therefore renders it an unreliable measure of performance. ‘Process’ considered qualitative aspects of how the group managed encountered problems, e.g. how they managed information, work distribution and decision making. To give an overview of the groups’ performances we will describe the groups’ processes and as mere illustrations we will report certain aspects of the groups’ outcomes.

2. Results: Descriptions of the three types of groups’ performances during the M/S Antwerpen simulation

Group type 1

The first type of group, in general, performed poorly in the first simulator session in regards to both process and outcome. Most of these groups lost the ship and barely saved any passengers on their first trips. Their work began directly with getting immersed in the simulation, literally just starting and going wherever events would take them, without any clearly communicated goal or strategy, thus becoming locked in the dynamics of events almost immediately (this initial approach provides information about the groups engagement
and the simulation’s face validity, see Dahlström et al., 2007). This was manifested in the group’s performance; they soon became overwhelmed by the amount of information that they received. At no time did they discuss whether their initial approach of handling the information, in the form of the printouts, was successful or even functioning at all. Actions were almost exclusively reactive in response to emerging problems, and were based on urgency, not priority. A minimal number of proactive actions were taken and tasks were not shared amongst group members outside their original role descriptions. The group did not monitor, discuss or change the group processes in any way, even when the workload was recognized within the group as becoming impossible to manage.

In the second session this type of groups did better regarding the outcome of the simulation: all groups still lost the ship, however, this time they managed to save the majority of the passengers and crew. On a process level, the improvements were even more evident. The group showed great improvements in general skills as communication, decision making and information sharing. Although this is partly expected due to the conclusions the participants drew from the first trip and the theoretical instructions (including topics as reactions in stressful situations, information management, group communication and decision making), they did display an enormous amount of transfer from the theoretical training to the manifested behavior in the simulation. Tasks were explicitly distributed and redistributed to manage workload within the group. A role with low workload took the role of moderator, who then was in charge of monitoring the group’s processes and dynamics. This gave the captain the opportunity to focus on other issues and had him maintain the overview. Measures to ensure effective information sharing were taken, such as attempts at establishing routines for regular briefings (successfully performed in the beginning) and many different forms of presenting the current situation of the ship, personnel, and problems to establish a shared mental model of the situation. In general, the groups did cope more effectively with problems and emergencies during the simulation; orders were no longer repeated and the group members had an improved overall awareness of the present situation of the ship in regards to location, staffing and the phase in which the emergency presided. They also followed up on more orders to ensure that they had intended effect (i.e. “effect control”) and were much more cautious when making assumptions when there was a lack of knowledge. The groups were more precautions and proactive during this trip, for example by checking that rescue equipment was operative.

In some instances these groups, being student pilots, lacked a general understanding of maritime concepts which made it challenging for them to extrapolate from available knowledge in order to invent creative solutions for solving problems. Clever alternatives were devised for general problems, for example, an effective notification was devised for passengers to tell them that as the general alarm was defective and that passengers should move to the life boats when the ship’s horn was sounded three times.

*Group type 2*

During the first trip, the second type of groups performed, in regards to final outcome, moderately better then the first group, however, most still lost the ship and the majority of the passengers and the crew. These groups devised creative solutions which were difficult for the first type of groups to come up with, since a basic understanding of ships is required for such solutions. As an example they used a stream anchor to navigate when the steering engine broke down. However, along side their creative problem solving capabilities, in comparison with the first type of group this type of group only performed marginally better concerning process aspects of the group work. Similarly to the previous groups, almost all of the second
groups’ actions were reactive; discussions and action were prompted exclusively by the
information that came out of the printer. Only a few proactive measures were taken. There
was hardly any analysis of the occurred events, revision of current situation and strategies,
and speculations on which potential tasks and risks could be ahead of them, instead idle time
was mostly used for personal conversations. There were also no structural solutions to deal
with the information overflow. The lack of protocols (and a division of labor) for example
resulted in making it everybody’s task to handle the printer output, adding to the confusion in
many situations.

After reflecting on the first trip and receiving the theoretical training, this type of group
performed very well in the second trip; saving the ship and its passengers. They were still
applying their creative problem solving skills, but now with the addition of proactive
thinking. They questioned their initial approaches to problems, and what became very
apparent during this trip was their suspicion in regards to potential problems. An example of
this was given as one participant did not accept a report from firemen which stated that a fire
was under control, and responded by sending in more firemen to ensure that everything was
under control. Examples of improved effect control were evident in the groups’ continuous
verification of the results of all their actions. There were numerous examples of proactive
actions and planning in regards to potential threats to the safety of the ship. During a large fire
all these groups sent people to higher decks, anticipating the spread of the fire. Their general
suspicion lead to them reacting in a thorough manner, that is quickly and forcefully; they sent
in a large amount of fire-fighters immediately instead of sending one man first to check what
was going on and thereby losing precious time (this being standard practice by all other types
of groups in the first session). They prioritized effectively by responding in this way to every
event that threatened safety and actively reduced the priority of less significant actions until a
normal status of operation was regained. They were also reluctant in making any sort of
assumptions which could not be sustained by facts. To the level of where they were having
expectations, they dealt very well with situations where expectations were not met,
dynamically adjusting their strategies. An example of this is their discovery, prior to any
emergency situation, that the instructions did not include specific references to muster stations
and the resulting creation of an alternative approach to evacuation.

These groups performed better than the first type of groups not only because they were more
cautious (the first group was very cautious as well), but mainly because they had contextual
knowledge that supported a more effective use of the general tools to manage situations more
effectively. They were more adapt in generating creative solutions based on extrapolations of
their nautical knowledge, e.g. reducing roll angle by course change and using auxiliary
engines proactively.

**Group type 3**

As this type of group consisted of participants with first hand practical experience in direct
relation to the roles in the simulation (the Captain, Chief Officer, engineers and Radio Officer
all had experience in these respective professional roles), this group initiated in the first trip a
steep hierarchy of command. This hierarchy seemed comparable to that of real vessels, with
the captain being clearly in charge of everything. This hierarchical role division and task
distribution worked effectively during this session as it provided structure to unclear
situations. All group members reported to the captain, this meant that the captain (and only
the captain) was in control of all goals and problems, and he gave orders accordingly. He then
relied on his crew to execute the given orders. The captain and rest of the group relied heavily
on many types of procedures that they would also apply in real-life situations, which proved
useful on most occasions. These procedures, drawn from practical experience, led to the participants checking a number of parameters and also proved useful as this supported a structured plan for management of emergencies. Such plans and clear task divisions were not made by the other groups on their first trips. However, the group did not fully integrate and act on information given to them, even though this issue was clearly stressed by the instructions they were given. This was manifested in many instances, and among them was one including a fire where the captain ordered the crew to go to their muster stations and only then discovered that these were not specified in the simulation. The group however was unable to solve this problem by inventing an alternative to the concept of muster stations; instead there were various repeated questions in the group regarding the use of muster stations, in addition to complaints to the facilitators that these were not there. Overall the procedures, often unspoken and simply assumed, did not always work as anticipated, because the scenario of the simulation did not match entirely with their expectations or with their previous experiences. As a result the group became lost and was not able to advance much beyond their silent consensus on how things “should be”. The group expected to be in possession of certain equipment and that this equipment would be in a functional condition. There were also expectations on the behavior of the crew which did not prove valid. The amount of expectations and inability to break out of them resulted in unanticipated failures and losses. In general their reliance on established procedures and direct real world experience led them to assume much more about the ship and the situation than the other groups did. The hierarchical team structure worked smoothly initially, however, this started to erode as soon as the workload increased; the captain became buried in information and tasks. For example, during one large fire the captain was in charge of the fire fighting and evacuation of passengers. This meant that only one of these two tasks being adequately managed. The final outcome was moderately successful because, based on his experience the captain decided at an early stage of the emergency that they were not going to manage the fire effectively and therefore he decided to evacuate.

Surprisingly, after the relatively good result of the first trip (keeping in mind the flaws of relying solely on procedures mentioned above) and subsequent training, this group showed little improvement in the second session. The group remained fully reliant on the roles and procedures that proved less than ideal during the first trip. In many aspects mistakes were identical to those committed during the first session. The group maintained a reactive approach towards the dynamics of events and showed little initiative to engage in information sharing or proactive performance. In addition there was a repetition of the lack of monitoring of group processes. The serial approach to problems, which was ineffective at times when the captain was preoccupied, was also maintained, leading to a standstill until the captain was free to approach the next problem. With the exception of maintenance of machinery, practically nothing was taken care of in parallel. An example of this was seen as the group was attempting to investigate potential water penetration, at that time a bomb threat was delivered, resulting in full focus on the bomb threat and they returned to the potential water penetration only when the bomb threat was over. By not approaching these two problems in parallel, for which there are more than enough resources available, the lower decks could have been flooded before the problems had been addressed. The group members stayed precisely within the boundaries of their roles and task descriptions (over relying on other “roles” to do “their own” tasks). It took a “hero”, one of the most experienced group members, the Chief Engineer (in real life and in the simulation), to break out of this structure. The Chief Engineer started questioning the captain’s actions and decisions during the second session. During an emergency he stepped out of the hierarchy to suggest to the captain that he should take charge of the evacuation as long as the captain was preoccupied with fire fighting. However, none of the other participants who had less demanding tasks did the same. Due to this initiative from
the Chief Engineer, the evacuation was taken care of effectively and again the outcome was relatively successful (although not as much as for the type 2 groups).

3. Discussion

3.1 Group type 1 & 2 vs. type 3: the limiting effect of procedural experience

Overall, the third type of group did relatively well in terms of group process and outcome. Despite that they did not perform optimally under increasing stress they did have a clear solution to the information flow problem and a clear distribution of tasks, as well as a systematic approach to most of the situations encountered. Most of this was based on reliance upon their knowledge of common procedures and normal management of a ship. However, as soon as events developed beyond the “design base” of their procedures, the work structure started to disintegrate rapidly. This did not occur in the other two types of groups – ones that were not as reliant on presupposed procedures (group type 1 and 2). These two groups used the general competencies and the knowledge they gained from the training in-between the two scenarios. Furthermore, in contrast to the other two types of groups, the third type of group very rigidly held onto their original hierarchical group structure, even after this proved ineffective. The other groups, after their first trip, were more able to recognize the changes in workload and in response adjusted their group structure accordingly; they applied a more dynamic group structure in order to deal with the different phases of escalation.

The third type of group had (or thought they had) a predefined procedure for every occasion, and therefore relied solely on procedural ways of dealing with events. Although procedures may prove effective to some degree, as shown by the relative success of the third group during their first trip, as soon as the specific conditions for a procedure did not precisely apply, control was lost. Interventions did not result in the expected outcome. As not all scenarios or situations can be foreseen, relying on procedural knowledge or procedures may not always suffice. However, trained crews tend react to emergency events with the use of pre-prescribed procedures even when these do not match the problem. In other words, for a group equipped with a hammer all problems will look like nails, and this is where the danger lies. During emergency situations especially individuals and groups need to be able to review their situation and assess whether their “standard” approach is appropriate. This process of distancing might then be particularly difficult for groups who are ingrained with predefined procedures, as in the example of the accident with Swiss Air 111 (Transportation Safety Board of Canada, 2003). In this way relying on procedural knowledge limits; the possibility for the positive impact of reviewing the situation, attempting other solutions than those pre-trained, and use of general problem solving skills to invent or improvise solutions to problems encountered. In short, relying on procedural knowledge can severely limit crews’ options to be resilient.

Moreover, armed with the knowledge that they have a procedure for every occasion, the experienced participants (group type 3) felt more secure than the inexperienced groups (group type 1 and 2). This false sense of security seemed to cause overconfidence. In the simulation this was manifested by the following behaviors: giving non-specific orders, not following up on orders and making unfounded assumptions in many situations. The most significant of these assumptions included ones regarding the specific knowledge of the crew on what they were required to do, and in regards to the availability and condition of equipment. This may work when in steady conditions; on a familiar ship with a familiar crew. However, this was not the case in this simulation and is rarely the case in the transportation sector, as crews do change frequently. This is also the case in many other safety related domains. (Participants of
the simulation were clearly instructed that they are new on the ship and that they should not assume anything.)

Despite long periods of operation during which nothing dangerous ever happens, it is of great importance that operators go to work and are prepared to expect the unexpected (Dekker, 2006). In the simulation sessions, in particular during the second trip, group types 1 and 2 showed caution regarding their situation; institutionalizing effect control, proactively checking the status of equipment (e.g. technical and rescue equipment) and simply contemplating what might go wrong next. Multiple groups of type 1, for example, checked whether all the life-boats were in good condition, whereas the experienced seamen assumed that they were, based on the claim that such matters are regularly tested (even though they admitted that there have been accidents at sea where rescue equipment has not been in functional condition). This summarizes exactly the issues that experienced crews face (group type 3); they assumed that things were as they should have been, that situations would proceed as expected or as ordered, and all of this in accordance with procedures. Experience makes people expect certain things regarding; quality of equipment, action sequences performed by crews, and crew reactions in emergencies, etc. Procedures create reliability, i.e. expected events. This is why experience and procedures risk limiting crew and staff vigilance. Procedures are in place to fight foreseeable problems, whereas caution and inspection can give rise to resilience. This can enable staff to question situations and actions and thus make them more resilient as operators (see Hollnagel, Woods, & Leveson, 2006).

A point of criticism may be that the experienced crew did not do as well as the other type of groups during the simulation because the simulation did not resemble the real world. This is however to some degree the main point of the simulation; in an emergency situation not everything will go according to plan, e.g. not everybody will report back and not all orders will be carried out (Dahlström, van Winsen, Dekker, & Nyce, 2007). But even under those circumstances groups need to stay functional and when procedures limit options they need to be able to find alternative ways to solve problems (and for a start recognize that not all problems can be anticipated and proceduralized).

3.2 Group 1 vs. group 2: Difference between general competencies, domain knowledge and procedural knowledge

If relying entirely on procedures is not appropriate, we can continue to argue that having only general competencies without having any domain specific knowledge (group type 1) can have the same effect. Maritime knowledge should facilitate the generation of solutions to the problems encountered in the simulation. The student pilots (group type 1) were not likely to come up with alternative solutions for navigation at sea. The second group type did have knowledge of maritime concepts in addition to the general competencies trained during the training program; even when the rudder engine broke down they explored actions in order to make the bow of the ship turn into the wind thereby reducing the dangerous roll angle of the ship. Furthermore they were not limited to following established procedures, even when there were explicit procedures available. It seems as if there is a benefit, even in a simulation constructed for training of general competencies, to be in possession of contextual competence from the domain. The constraints of the situations encountered might not allow people to invent the wheel every time a problem arises.

On a process level, it proved easier for groups with a certain level of domain knowledge (group type 2), as compared to groups with only general competencies (group type 1), to cooperate. The reason for this lies in the more appropriate match between their common
knowledge and the situation. The type 1 groups improved their general competencies between
the two trips and therefore performed more effectively during the second trip. Despite this
improvement, it proved hard for these groups to apply their skills in an unfamiliar setting.
Moreover, for participants without previous domain-specific knowledge, it proved impossible
to step out of their roles, in order to assist or even overrule other roles. The reason for this lays
in the role-specific information being the only information on which to base their behavior.
With no alternative securities in these unfamiliar settings they remain ‘too loyal’ to these
roles.

There is a fine line that distinguishes the second group’s (theoretical) domain knowledge,
which enabled them to come up with more creative solutions, from the first group who lacked
maritime concepts, and the third group’s procedural experience which seemed to limit them to
following procedures. The first group had no tools with which to deal with a maritime
emergency, they had to rely solely on general knowledge and competencies. The third group
relied on procedures, tools and procedures they normally used. The second group on the other
hand was able to look into the toolbox and select an appropriate (but not necessarily
proscribed) tool for the situation. This enabled them to fit strategy and resources to the
situation and not to function as merely bricoleurs and as such employing a “strategic
flexibility”.

References

Science, 37, 109-126.

Reports. Proceedings of the 12th International Symposium on Aviation Psychology. Dayton,
OH: Wright State University Press.

simulation. Submitted to the Theoretical issues in ergonomics science.

Dekker, S.W.A. (2001). Follow the procedure or survive. Human Factors and Aerospace
Safety, 1, 381-385.


Company.

Rethinking pilot error and the causes of airline accidents. Aldershot, UK: Ashgate Publishing
Company.


