Teams organising ‘work as done’: resilience, repetition, and expertise

Joseph Paulus Havinga, BSc, MSc

Supervisors:

Prof Sidney W. A. Dekker

Dr Andrew Rae

Dr Hamish McLean

Safety Science Innovation Lab

School of Humanities, Languages and Social Science

Arts, Education and Law

Griffith University

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

Date of submission: October 2017

Abstract

This thesis *Teams organising 'work as done': resilience, repetition, and expertise* presents a series of studies and research papers on how teams organise their work for safety, efficiency and a host of other goals. In a sense, it aims to inform on improving how to organise teams and organise for teams, but this is expected to arise from a better understanding of 'work-as-done', rather than imposing a 'work-as-imagined'. A broader view of initiatives across a number of industries is part of this thesis, however, the major empirical focus is on 'work-as-done'. Empirically, the research for this thesis was driven by the work of semi-autonomous blast crews who operate in open-pit mines and quarries, as well as electrical line crews deployed across southeast Queensland. The thesis also took a broader view that makes its findings applicable to other teams as well. The thesis is organised around four papers, which form a literature review, a methodological reflection, an empirical investigation, and a theoretical paper extending the empirical work.

The thesis began by investigating the-state-of-the-art way to organising teams, by reviewing Crew Resource Management (CRM) literature across industries. The review discovered that there was very little evidence to support that CRM achieve their goals of improving safety and efficiency. In addition, there was no unified conceptualisation of how CRM is supposed to reach those goals. This diversity in conceptualisation made it impossible to further investigate whether CRM is working as intended, as the questions to investigate differed per conceptualisation. This review showed there is little support for teaching a state-of-the-art way of organising to people, nor a clear idea of how teams operate.

The next study in the thesis explored the methodological considerations for investigating everyday work for safety purposes. By considering the possibilities and different schools of thought in safety, I reflected on the design of investigations of everyday work for safety. I found a general tension between trying to capture and learn from the everyday, and the goal of investigating something that is known to be relevant to safety. What is known the be relevant can be hard to fit with the reality of everyday work and could steer away from

exploring new areas, while with the everyday there is the uncertainty of whether something is relevant for safety.

These reflections led to the design of the methodology for the third study. In this study, I applied problem-oriented ethnography to investigate how blast crews and linesmen organise routine work. From my results, I explained how individual decisions give rise to general repeated patterns of work at a team level. I found that operators have a stable problem understanding of and repeating patterns in how they approach a task. This repetition is not a replication of actions, but in the way teams divide and structure a larger task into smaller goals — their solution structure. As team members work together, their problem understanding and solution structure converges, which leads to smooth and coordinated work process of a team. The problem understanding and solution structure capture part of the expertise of crew members and allow them to perceive meaningful signs in their environment. The repeating patterns with which teams complete tasks makes them more sensitive to their operations, as deviations will stand out. In addition, crew members could interact adaptively because of it, as they could use intentional deviations to alert each other.

Building on the finding that even for routine work, team members rely on expertise to be adaptive, in the fourth study I explored the relationship between expertise and resilience, and how to manage expertise on the level of frontline work, teams and management, and systems to enhance resilience.

The conclusions from this thesis explain how teams organise routine work resiliently. Unlike the work on High Reliability Organisations and Resilience Engineering, my conclusions stress how repetition helps teams collaborate and cope with complexity, while still allowing for the needed adaptations. My conclusions challenge the distinction between heedful and routine action from the High Reliability Organisations literature. The thesis extends theories from Cognitive Systems Engineering on how team processes can shape the goals of individuals. My conclusions contradict organisational routine theories on what repeats in a routine and extend theories of organisation routines as to how actions in a routine link together. In addition, I made methodological contributions to investigating of everyday work for safety and the study of changing routines. In terms of practical application, the thesis makes suggestions on team member composition and rotation, as well as on the specification of procedures.

Statement of originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Joseph Paulus Havinga

October 2017

Table of Contents

List of figures

List of Tables

List of acronyms

ATC	Air Traffic Control
CMAQ	Controller Resource Management Attitude Questionaire
CRM	Crew Resource Management
FMAQ	Flight Management Attitude Questionnaire
HRO	High Reliability Organisations
KLM	Koninklijke Luchtvaart Maatschappij
MAR	Maritime industry
NDM	Naturalistic Decision Making
NOTECHS	Non-TECHnical Skills
NPI	Nuclear Power Industry
NASA	National Aeronautics and Space Administration
O&G	Oil and Gas Industry
RE	Resilience Engineering

Acknowledgements

I would like to thank my supervisory team, Professor Sidney Dekker, Dr Andrew Rae, and Dr Hamish Mclean. Sidney, thank you for setting up this PhD project and choosing me as the person for it. Thank you for your energy, brilliant brain, and thought-provoking way of thinking. Drew, thank you for your dedication, and guidance, as well as your clarity of thought, and open mind. Thank you both for providing the freedom to take charge of this project and helping me grow and learn in the process. I would like to thank Griffith University, the school of humanities, the Higher Degree Research Student Centre, the Postgraduate student association for the services they provide that have supported me, including their workshops and editing assistance.

My thanks also go out to Orica, which has provided the research opportunity, as well as the means for me to do this research with a scholarship. David and Graham, thank you for setting up this project. Scott, Nicki, and Ray, thank you for helping to organise the research. All employees that have participated in my research and have welcomed me on their site and in their team, thank you.

Also, I want to thank everybody I have met through the Safety Science Innovation Lab and made it into a great time there. David and Roel for guiding me when I was new. Daniel, thank you for having the vision to build Art of Work and taking me on. Guido, thank you for our great discussions. Dave, thank you, for sharing your business experiences and sharp reflections. Charan, thank you, for your great company and curious mind. Jonathan, thank you for all of the things mentioned above, as well as the coffee in the morning.

Finally, I want to thank my parents, my brother, friends, and of course partner. While less directly involved with this work, the knowledge that you would be there for me if I would need it, have helped me complete this thesis.

Without any of you this thesis would not be here in the shape it is.

Acknowledgement of Papers included in this Thesis

Section 9.1 of the Griffith University Code for the Responsible Conduct of Research ("Criteria for Authorship"), in accordance with Section 5 of the Australian Code for the Responsible Conduct of Research, states:

To be named as an author, a researcher must have made a substantial scholarly contribution to the creative or scholarly work that constitutes the research output, and be able to take public responsibility for at least that part of the work they contributed. Attribution of authorship depends to some extent on the discipline and publisher policies, but in all cases, authorship must be based on substantial contributions in a combination of one or more of:

- *conception and design of the research project*
- *analysis and interpretation of research data*
- *drafting or making significant parts of the creative or scholarly work or critically revising it so as to contribute significantly to the final output.*

Section 9.3 of the Griffith University Code ("Responsibilities of Researchers"), in accordance with Section 5 of the Australian Code, states:

Researchers are expected to:

- *offer authorship to all people, including research trainees, who meet the criteria for authorship listed above, but only those people.*
- *accept or decline offers of authorship promptly in writing.*
- *Include in the list of authors only those who have accepted authorship*
- *Appoint one author to be the executive author to record authorship and manage correspondence about the work with the publisher and other interested parties.*
- *Acknowledge all those who have contributed to the research, facilities or materials but who do not qualify as authors, such as research assistants, technical staff, and advisors on cultural or community knowledge. Obtain written consent to name individuals.*

Included in this thesis are papers in *Chapters 2, 3, 4, and 5, which* are co-authored with other researchers. My contribution to each co-authored paper is outlined at the front of the relevant chapter. The bibliographic details or status for these papers including <u>all</u> authors, are:

Chapter 2: (paper in press)

Havinga, J., de Boer, R. J., Rae, A., Dekker, S. W. A. (2017) How did CRM take-off outside of the cockpit? A systematic review of how Crew Resource Management training is conceptualised and evaluated for non-pilots. *Manuscript accepted for publication by Safety.*

Chapter 3: (published paper)

Havinga, J., Dekker, S., & Rae, A. (2017). Everyday work investigations for safety. *Theoretical Issues in Ergonomics Science*. http://doi.org/10.1080/1463922X.2017.1356394

Chapter 4: (paper under review)

Havinga, J., Dekker, S., & Rae, A. (2017). Routine variation: How routines facilitate adaptation. Manuscript submitted to *Cognition Technology & Work*

Chapter 5: (book chapter in press)

Havinga, J., Bergström, J., Dekker, S., Rae, A. (2019) In Ward, P., Schraagen, J. M., Gore, J., Roth, E. (Eds.) The Oxford Handbook of Expertise: Research and Application. Oxford, Oxford University Press

(Signed) ______________________________ (Date)____________

Joseph Paulus Havinga

(Countersigned) ______________________________ (Date)______________

Supervisor: Sidney Dekker

(Countersigned) ______________________________ (Date)______________

Supervisor: Andrew Rae

Chapter 1: **General introduction**

This thesis investigates how teams organise their work for safety, efficiency and a host of other goals. It is organised around four published, accepted, or submitted pieces of work:

- A literature review
- A methodological paper
- An empirical investigation leading to new theory
- A theoretical paper extending the empirical work

The concept of organising for safety has become popular with the school of High Reliability Organisations (HRO). The distinction between "work-as-done" versus "work as imagined" has gained almost canonical status in work/organisation studies, Cognitive Systems Engineering (CSE), and Resilience Engineering (RE). This thesis focuses on 'work as done'. The only "work-as-imagined" addressed in the thesis is the work as imagined by the practitioners doing the work themselves. This opposed to how "work-as-imagined" is generally used, namely as imagined by people managing and planning the work. The literature on organisational routines is used to help structure and then build an understanding of how workers organise their actual work around their own image of work.

The main aim of the thesis is to investigate how teams organise for safety. Empirically, it draws on the everyday work of blast crews, and to a lesser extent, electrical line crews. This search is not driven by the idea that blast crews in mining are unsafe in their work, or even that they are in any particular need of improvement. The opposite actually seemed true. The company involved had suffered very few serious accidents with its blast crews. Nonetheless, the rate of serious accidents had stopped decreasing and had reached a plateau, suggesting that current strategies were not working to reduce accidents anymore. Critical safety innovations are not expected to come from new technologies. Instead, hopes were pinned on a better understanding of the complexity, social construction, and organisation of work as it actually happens in the quarries and mines.

This thesis builds on four bodies of research. The first is Resilience Engineering, as Resilience Engineering provides the view of safety used in this thesis. The second is High Reliability Organisation, because this thesis looks at how to organise for safety. The third body of research is Crew Resource Management training, as CRM aims to teach a state-of-the-art way on how to organise teams for safety. The fourth being organisational routines, as it fits with the scope of how everyday work is organised.

1.1 Resilience Engineering

David Woods (2003) introduced Resilience Engineering (RE) in a testimony on the future of the National Aeronautics and Space Administration (NASA) to a U.S. Senate Committee. After a symposium on Resilience Engineering the groundwork was laid for this new discipline (Hollnagel, Woods, & Leveson, 2006). The ideas of Resilience Engineering grew out the application of Cognitive System Engineering ideas to the field of safety. Cognitive Systems Engineering is about "how humans can cope with and master the complexity of processes and technological environments" (Hollnagel & Woods, 2005, p.1). Coming from studying coping and mastery, it is not surprising that Resilience Engineering emphasises on how success is achieved.

This emphasis on success sets Resilience Engineering apart from other safety paradigms that focus on avoiding accidents. Avoiding accidents often leads to limiting variations within a system and constraining individual agents, so an organisation cannot follow a path associated with accidents. Resilience Engineering, on the other hand, studies on how systems achieve success, cope with stress and complexity, and build adaptive capacity (Bergström, van Winsen, & Henriqson, 2015). The argument for making systems more adaptable relies on the law of requisite variety, which says systems need to be able to vary their approach, to deal with varying circumstances (Ashby, 1957). By seeing variation as necessary, Resilience Engineering's goal changes from trying to constrain systems and trying to keep the system as close to the original design as possible, to trying to enable systems to adapt to disturbances where needed, even beyond the point of what a system is specified to do and handle (Hollnagel et al., 2006).

This change in view on variation, leads to a different appreciation of humans in a system, because humans are considered a source of variation. In traditional safety paradigms, that people can vary their behaviour makes people risk factors that can cause deviations from the system as designed. Seeing people as a risk factor leads to controlling and constraining the actions of people, for example with rules and procedures. Behaviour based safety (Krause, 1997), for example, focuses on making all practitioners act the same by having them follow procedures and avoiding any deviations that are considered unsafe. For Resilience Engineering, that people can vary their actions makes them desirable, as this enables the system to deal with unexpected things. Instead, the focus is of Resilience Engineering is on what helps people to make successful adaptions, such as their expertise on a subject, information available to them, or resources they need.

Inspired by Resilience Engineering, the aim of this thesis is to look for things that make teams more capable of handling variations and finding ways to deal with these variations, rather than limiting the options how teams can or should organise their work. This is especially relevant for blast crews, as the areas in a mine are continually changing. In addition, different blast crews work in different mines, with different conditions, making it less likely there is one "best" way to perform a task, or actions to always avoid. Blast crews will always and continuously need to adapt to their circumstances.

1.2 High reliability Organisations

The High Reliability Organisation (HRO) school began with a group of researchers from the University of California, the so-called Berkeley group (Bourrier, 2011), who studied "the design and management of hazardous organisations that achieve extremely high levels of reliable and safe operations"

(Roberts, 1989, p.111). The researchers of this group included Todd La Porte, Karlene Roberts, and Gene Rochlin.

The group interested in this topic continued to grow as ideas developed and as it broadened its horizon. Karl Weick, an organisational theorist, became involved with the research group and contributed with ideas of his own. Through Roberts in 1995, another group of safety scientists at the California State University in San Bernadino started to work on HRO ideas. The San Bernadino group expanded the HRO work to business failures, rather than just accidents, and further developed the HRO ideas. Over time, the HRO work grew into a set of common principles and a similar way of looking at highly complex and hazardous organisations.

The early HRO work focused on describing how organising happened, and did not recommend on how to organise. There were, however, returning characteristics found in almost of the studied HROs (La Porte, 1996; K. H. Roberts, 1990; Karlene H. Roberts, 1989):

- Organisational structure and rules: HROs were marked by flexibility and redundancy in their structure and rules. These appeared in the form of parallel and overlapping processes, backups, skills redundancy, and cross-checks. Redundancy formed a way to recover from errors and ensure reliability. In addition, HROs could quickly switch from a hierarchic organisation during routine operations to patterns deferring to experience in demanding situations.
- Operational decision-making and communication: The flexible adaption of organisational structures changed the way decisions were made and information was communicated. If circumstances required, decision-making was passed to the people with the most experience and the necessary information at hand for the problem. In addition, HROs often had channels of communication to constantly to inform others, to establish and maintain a “larger picture” and to keep teams integrated.
- The “culture” of high reliability: Within HROs people were closely engaged with understanding the current operations and constantly looking for things that could go wrong. Compared to other organisations, in HRO’s people were very

willing to report errors and potential failures. People were not blamed for reporting errors and could even be rewarded for this.

- The adaption to technology: Within HROs staff appeared to have paradoxical attitudes towards new technology. They were remarkably adept at integrating new technologies into their operations, while also resisting the introduction of new technologies.

Later HRO work changed from this descriptive stance and began issuing instructions and recommendations on how to organise for reliability. Weick and Sutcliffe (2007) introduced the idea of mindfulness as the central element of HROs. They defined mindfulness as a “a rich awareness of discriminatory detail” (p. 30). Mindfulness increases the capability to detect and handle unexpected events, which leads to reliability in outcomes, even in varying conditions. They mention five processes that create mindfulness;

Preoccupation with failure: This process reflects the ongoing focus on failure of staff within a successful HRO. Not only accidents, but also near failures and near-misses, are seen as a sign of danger. People are sensitive to ‘weak’ signals, signals that when they happen do not have a direct link to danger, but can be read as signs there are vulnerabilities or disturbances in the system. In addition, HRO staff does not see past successes as a sign of competence and that they can relax. They remain vigilant and stay alert for how things can go wrong.

Reluctance to simplify interpretations: Processes within HROs are aimed at preventing simplifications, because simplifications can lead to negligence of data, reduction of precautions, and increased likelihood of undesired surprises. This takes shape in renewal and revision, job rotation, and cross-checks.

Sensitivity to operations: HROs put effort into establishing and maintaining a big picture of operations. Because all individuals only have access to part of the information, maintaining a big picture is a shared effort. HROs achieve this through collective story-building, shared mental representations, and knowledge of system parameters.

Commitment to resilience: Besides preventing surprises and disturbances before accidents happen, in HROs people also plan to be able to respond to disturbances when they occur. They prepare for the unexpected to happen.

Deference to expertise: While HROs are normally structured in a hierarchical way, they can quickly adjust and give decision authority to people with the most expertise. In some cases, this is referred to as "under specification of structures".

As HRO's are mindful in their operations, they do not simply follow routines where they replicate what they have done before. They constantly adjust what their actions to fit with the current conditions and the pursuit of larger goals. This opposes the idea that organisations achieve reliability by developing and following highly standardised routines (Weick, Sutcliffe, & Obstfeld, 1999). HRO literature multiple times relates routine action as something mindless, replicating to what has been done before. Routine actions oppose mindful actions, or heedful actions, where people consider their actions as contributions to the larger whole, rather than as solitary acts (Weick & Roberts, 1993; Weick & Sutcliffe, 2007).

HRO research does not take accidents, or their causes, as a dependent variable. Instead, HRO looks for operations that are nearly error-free. It is interested in how organisations create high reliability and safety levels (Roberts, 1990). With this approach, HRO is like Resilience Engineering, and more interested in what contributes to success, rather than what causes accidents. Also in agreement with Resilience Engineering, HRO says that adaptations constantly occur and are necessary for organisations to reach reliable outcomes. Flexibility in structure and redundancy in capacity was needed for organisations to adapt to unexpected changes. However, the HRO research focuses on studying groups that are successful and tries to find out what they do differently. Resilience Engineering, on the other hand, tends to look more at what is, and what systems do and to they adapt (Haavik, Antonsen, Rosness, & Hale, 2016).

Blast crews have a good safety record. In that regard, they fit nicely as an object of study as an HRO. Blast crews can most likely provide an example for how to organise teams, and could reveal differences with the more typical high-tech domains investigated in HRO's.

1.3 Crew Resource Management

HRO is of course not the only body of research available on how to organise for safety. Crew resource Management (CRM) training has long been used to teach cockpit crews on how they should work together to improve safety and efficiency. The first training programs in aviation that addressed interpersonal and social aspects of crew behaviour emerged in the 1970's. In the Netherlands, the Koninklijke Luchtvaart Maatschapppij (KLM) based their training on the Edwards' SHELL model and trans-cockpit, most of this training focussed on the captain. In the late 1970's a number of aviation accidents occurred where poor communication and breakdowns of teamwork in the crew were mentioned as a cause. In response to this, NASA organised a workshop in 1979 to discuss the theoretical and operational concerns of flights from a human factors point of view (Cooper, White, & Lauber, 1979). Many academic and government researchers, along with aviation managers, were present and left with the ambition to develop formal training on the subject. This lead to multiple early CRM courses in the 1980's (Kanki, Helmreich, & Anca, 2010; Salas, Wilson, Burke, Wightman, & Howse, 2006).

Lauber (1984) initially defined CRM as "using all available resources—information, equipment, and people—to achieve safe and efficient flight operations" (p. 20). At this point, the CRM still stood for cockpit resource management, but cockpit changed to crew as the scope of CRM training broadened to include interaction with automation and relations with the rest of the flight crew (Kanki et al., 2010). However, other definitions have been suggested as well, that emphasise the improving teamwork (Salas, Wilson, Burke, Wightman, et al., 2006) or human error reduction (Helmreich, Klinect, & Wilhelm, 1999). CRM can be regarded as teaching crews the state-of-the-art way of organising teams for safety, as it teaches crews how to best coordinate with each other (Salas et al., 2008).

The earlier CRM training programs were met with some resistance, but over time they became well-accepted and so spread across airlines in all countries. Many people have now claimed that CRM is one of the greatest successes of aviation human factors (Kanki et al., 2010; Maurino & Murray, 2009; Salas, Wilson, Burke, Wightman, et al., 2006). Naturally, other industries have followed this example after this claimed success (Kanki et al., 2010). However, while CRM has been claimed to be a great success, the found effects of CRM during evaluative studies are limited (Edkins, 2002; Salas, Wilson, Burke, & Wightman, 2006). If it was because of evaluation results, this raises the question on what basis CRM training was changed over time and what was it that made people claim CRM was such a success.

I have found no indications of any existing CRM programs for mining personnel, let alone for blast crews. This meant the academic literature lacked specific suggestions on how to organise blast crews. However, because CRM has been adopted and adapted to other industries, there might lessons to be learned from how CRM has been adopted and adapted other sectors.

1.4 Organisational routines

The term organisational routines was introduced by Nelson and Winter (1985), but resembled older observations about how organisations work (Cyert & March, 1963; March & Simon, 1958). These were observations that what organisations do is rarely the product of calculating the best response, but instead a process of matching known procedures or scripts to the situation. Actions in an organisational process reflect a routine way of doing things, rather than the calculated optimal path. Routines were seen as a way organisations satisfice because of their bounded rationality.

From this view, organisational routines were understood as something “mindless”, related to low-skill work and inertia. This view is similar to how HRO links routine action to heedless actions that are not mindful of the big picture and not sensitive to the state of operations (Weick & Roberts, 1993). This view, however, begun to shift, when empirical studies found that variations of routines were the norm, not the exception, and that change was an inherent feature of many routines (Feldman, 2000; Pentland & Rueter, 1994). Organisational routines are now seen as both a cause for stability as well as change (Feldman & Pentland, 2003; Pentland & Hærem, 2015). Overtime, organisational routines have been used to study many types of work that are not considered mindless or low skill work at all, such as medical work, engineering, legal work, and research & development jobs (Pentland & Hærem, 2015).

In this thesis, I use the definition and conceptualisation of Feldman and Pentland (2003), who defined organisational routines as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (p.95). This is currently the most used definition in academia (Pentland & Hærem, 2015). "Repetitive" here means that the behaviour pattern is shown more than once. It needs to be a "recognizable pattern"; otherwise it cannot be determined whether there is a pattern let alone a routine. It concerns "interdependent actions", the actions in a routine build on the previous actions. For a pattern of actions to be a routine, it needs to be carried out by "multiple actors". If only one actor is involved, it is considered a habit.

The definition proposed by Feldman and Pentland (2003) was not innovative and was build on the 'core' of other definitions that were used before it. They did introduce a new view on the ontology of organisational routines. They proposed a duality of routines in an *ostensive* and *performative* aspect that influence each other and make up a routine together. This ontology allowed them to resolve the conflict of the inertia of routines versus the variability and change perceived in many empirical studies (Feldman & Pentland, 2003).

The ostensive aspect of a routine is the ideal image of a routine or a summary of how things are usually done. The ostensive aspect is what a practitioner, or researcher, recognises as the general repeated pattern. The ostensive is, however, not a single unified picture of the world. Different people can create a different image of the same activity. Each practitioner's image is shaped by their role in it. The ostensive aspect is abstract and does not include specifics, because it needs to be applicable in different settings. The ostensive is not the same as written rules or procedures, although those can be based on the ostensive. Rules and procedures are artefacts and require an interpretation before people can enact them (Feldman & Pentland, 2008; Reynaud, 2005; Wright, 2003).

The performative aspect is how people enact a routine. The performative includes the people who do it, the time, and the context in which the actions take place. The performative aspect is always improvisational and will contain deviations from the ostensive abstract image.

The ostensive aspect influences the performative aspect and the performative aspect influences the ostensive aspect of the routine. The ostensive aspect guides practitioners when they are performing a routine on what to do. The ostensive aspect allows practitioners to account for what they are doing and helps them make sense of what is going on. Finally, the ostensive aspect enables practitioners to refer to the performative, and therefore plan activity. The reality is too messy, too complex, and too incomprehensible to describe, while a summary or simplified ostensive image allows them to talk about it.

The performative aspect creates the ostensive aspect as practice happens, and maintains it through repetition. Through the variations when routines are performed, the performative aspect modifies the ostensive aspect.

By including making routines consist of both the ostensive and the performative, Feldman and Pentland (2003) could explain both the variation and stability observed in routines.

Organisational routines need to be understood in the context in which they take place. A “good” routine in one context can become meaningless and can even be related to poor performance when moved to another setting. This is why there is no such thing as a universal best practice, but only a best practice in context. Routines have been transferred to different contexts with success, but often not smoothly. Implementers of a routine often found it not clear what were the essential elements of a routine, that routines did not fit with a new context, or that some parts of a routine relied on tacit knowledge that was not transferred (Becker, 2004; Szulanski, 2001).

Routines have been said to embody multiple things and to serve various functions. Routines have been suggested to embody the knowledge of a group on how to complete a task. This shared expectation of a group reduces the need for explicit coordination (Nelson & Winter, 1985). When groups are well established and familiar with routines, the actions in a routine can coalesce, leading to a smooth following sequence of actions, basically without any explicit coordination (Turner & Rindova, 2012). Routines can reflect as a truce between the goals of multiple agents (Becker, Lazaric, Nelson, & Winter, 2005). In addition, routines are seen both as a source of stability and change (Feldman, 2000; Yi, Becker, & Yi, 2016). This variety of views on routines is related to the wide range of disciplines that use the concept.

Recognised challenges within this literature include the lack of a clear explanation of how actions in a routine link together (Pentland & Hærem, 2015) and that the line between variation within a routine and a change of a routine is difficult to identify (Feldman, Pentland, Adderio, & Lazaric, 2016).

As I will allude to in Chapter 3 and describe in-depth in Chapter 4, the way blast crews organise their work revolves around having a routine method of completing a job. As suggested in the organisational routine literature, blast crews only rely on planning and explicit coordination to a limited degree. In addition, organisational routines studies put the focus on patterns of activity, making it more agnostic to the actor completing an action. This agnostic stance makes it suited for describing blast crews at work, because there is little set task distribution and a lot of it arises as work is being completed and operators can quickly step in for each other.

1.5 Overview of thesis and research question

Our main aim of the thesis is to investigate how teams organise for safety, efficiency, and other concerns, with the intention to inform for the practical concerns how to better organise teams or how to better organise for teams. I approach safety as being able to succeed in varying conditions, as found in Resilience Engineering. To answer this, I look at what is prescribed on how to organise teams, how to study teams at work, how teams organise routine work, and explore how expertise links to resilience, as expertise was found to play an important role. For the empirical work, I study blast crews, and to a lesser degree, linesmen. An overview of the contribution to the research aims and research questions per question can be found Table 1, which I explain in-depth in the rest of this section. After the four chapters mentioned in

Table 1, the thesis recapitulates its contributions in chapter 6, the general discussion.

Table 1 Research questions and contribution per chapter

Chapter	Contribution to research aim	Research Question
2	Literature review on prescribed ways of organising teams	How has CRM been adopted across industries?
3	Methodology	How to investigate everyday work for safety purposes?
4	Empirical study of how teams organise	How do teams organise routine work?
5	Extension of results through review	How to manage expertise to enhance resilience?

In chapter 2, I investigate the state-of-the-art way of organising teams in different industries, with the question "How has CRM been adopted in other industries?". This review of the literature initially aimed to find how CRM was adapted to fit industries and to identify general patterns from this. However, while conducting this initial review, I found that the supposed adaptation of CRM to different types of teams and industries was hardly discussed in the literature. There seemed to be differences between papers, but there were few explicit statements why things were done, or how lessons drawn from other CRM research were taken into account. For the most part, I found no trace in the literature on the adaptation process, this despite a thorough and extensive literature review. On top of that, there few CRM training programs with soundly evaluated. So instead, the review focused on the steps preceding adaptation, that is, how CRM was conceptualised and evaluated during and after its adoption in different industries. With this, I hoped to assess whether CRM was a uniform enough intervention and had enough evaluation to support drawing more general conclusions from.

In chapter 3, I investigate how to study how teams organise their work in relation to safety, by reflecting on the question "How to investigate everyday work for safety purposes?". During this reflection, I make the consideration that goes into such an investigation explicit and relate these to different schools of thought in the safety literature. Using some initial observations made from blast crews at work, I explore how these questions can play out in practice. This chapter provides the methodological considerations of the empirical study in Chapter 4.

In chapter 4, I address "How do teams organise routine work?", through a study on blast crews and linesmen. The focus on routine work is driven by my initial observations on blast crews. These observations suggested that blast crews—despite being adaptive—also relied pretty heavily on routines, or at least on repetitive patterns to organise their work. This was seemingly at odds with the HRO and Resilience Engineering work, which focus on adaptations and variation. The organisational routines literature, on the other hand, suggests that relying on repeating patterns is common in all industries. The tension between the literature groups and my observations make it a topic worth investigating.

In chapter 5, I extend the findings from Chapter 4. I review the literature to answer the question "How to manage expertise to enhance resilience?". This based on the finding in Chapter 4 that even for routine operations people rely on expertise, and it is expertise that enables seeking alternative paths when conditions change. I look at the relation between expertise and resilience on the level of frontline workers, teams and management, and systems, as well as how to manage expertise on these levels.

Chapter 6 is the general discussion, and it recapitulates the contributions of the thesis to the understanding of organising for safety, to the various literatures it was inspired by, and to the understanding of practices by blast crews and similar teams in other organisations. The studies are summarised for their contribution to the main research aim. In addition to the contributions to the main research aim, I discuss the scientific contributions in terms of theory and method. The chapter ends with considering the practical applications of the research.

With the exception of the general discussion, each of these chapters is set up of as a stand-alone paper. Supplementary to the stand-alone paper, each chapter starts with a section that explains the rationale for the study and ends with a section that explains the contribution of the study to the research aims. The thesis as a whole, reflect my choices, but the separate studies are the product of collaborations with other researchers and also the reflect effort put in by these researchers. At the start of each chapter, my contribution to each paper is summarised and all co-authors are acknowledged.

Chapter 2: **Prescriptive ways to organise teams**

2.1 Statement of contribution of Co-Authored paper

This chapter includes a co-authored submitted paper. The bibliographic status of the co-authored paper, including all authors, is:

Havinga, J., de Boer, R. J., Rae, A., Dekker, S. W. A. (2017) How did CRM take-off outside of the cockpit? A systematic review of how Crew Resource Management training is conceptualised and evaluated for non-pilots. *Manuscript accepted for publication by Safety.*

My contribution to the paper involved:

The development of the research questions; The method design, with normal PhD supervision from Prof Dekker and Dr Rae; The preliminary search, review, and analysis of the literature; The creation of the first draft; The refinement of categories and search criteria in discussion with Dr Rae and Dr de Boer; The final search, review and analysis of the literature; The second draft; The incorporation of feedback from Prof Dekker, Dr Rae, and Dr de Boer.

(Signed) ______________________ (Date)__________

Joseph Paulus Havinga (Corresponding Author)

(Countersigned) ______________________ (Date)__________

Dr Andrew Rae (Principal supervisor and co-author)

(Countersigned) ______________________ (Date)__________

Prof Sidney Dekker (Principal supervisor and co-author)

2.2 Rationale for the study

Originally, the plan was to analyse how Crew Resource Management (CRM) programs changed across industries and to assess the effectiveness of different types of training programs in different industries. This approach was inspired by a few in-depth task analyses of the work in different industries. These studies pointed to the relevance of features such as the stability of team membership and the diversity of disciplines within teams. Comparison of these types of analyses across industries could identify general patterns that are relevant when training teams, as well as how to adjust training to different teams and conditions. Such a review could have served as a starting point for other industries trying to adopt CRM. Even more, such a review could point to underlying patterns of how teams work. These patterns could, in turn, be tested in further research or at least guide the focus of subsequent research.

However, during the initial review, some assumptions preceding these questions came to light. I expected CRM to be unified and that there would be clear reasons for adjustments between industries. However, it turned out that this was not the case, as I ran into conflicting statements about CRM. For example, some people described CRM as teaching something that is largely based on common sense, things everyone has experience with, and that the training is merely a facilitator of familiar material (Fonne & Fredriksen, 1995). Whereas others said CRM was about teaching new skills (Hackman, 1993), and that there was a large variation between what trainees had learned about it thus far (Gaddy & Wachtel, 1992). By itself, this was not an immediate problem, as diversity in training design, content, and philosophy, can be helpful to determine what CRM programs are and are not effective, as it creates an opportunity to compare CRM programs. However, while studying the literature, I also noticed there were few arguments or cases made as to why one type of training was better than another. In addition, it was often not clear what effect CRM programs had, as methodological sound evaluation of a training program was rare, making it impossible to investigate the effect of the differences between CRM programs.

Therefore, instead of looking at how CRM programs had been adapted to different industries, I instead decided to examine the assumptions that had guided me before and I investigated how CRM had been adopted in different industries, with a particular focus on how CRM was conceptualised and evaluated.

How did CRM take-off outside of the cockpit? A systematic review of how Crew Resource Management training is conceptualised and evaluated for non-pilots

2.3 Abstract

Crew Resource Management (CRM) training for flight crews is widespread and has been credited with improving aviation safety. As other industries have adopted CRM, they have interpreted CRM in different ways. We sought to understand how industries have adopted CRM, regarding its conceptualisation and evaluation. For this, we conducted a systematic review of CRM studies in the maritime, nuclear power, oil & gas and air traffic control industries. We searched three electronic databases (Web of Science, Science Direct, and Scopus) and CRM reviews for papers. We analysed these papers on their goals, scope, levers of change, and evaluation. To synthesise, we compared the analysis results across industries. We found that most CRM programs had the broad goals of improving safety and efficiency. However, there were differences in the scope and levers of change between programs, both within and between industries. Most evaluative studies suffered from methodological weaknesses, and the evaluation did not align with how studies conceptualise CRM. These results challenge the assumption that there is a clear link between CRM training and enhanced safety in the analysed industries. Future CRM research needs to provide a clear conceptualisation — how CRM is expected to improve safety — and select evaluation measures consistent with this.

2.4 Introduction

The development of CRM for flight crews has been repeatedly called a great success story (Kanki et al., 2010; Maurino & Murray, 2009) and sometimes even “one of the greatest successes of aviation and the human factors/ergonomics field”(Salas, Wilson, Burke, Wightman, et al., 2006, p.36). While it initially was met with some resistance, CRM is now well accepted and has spread to all countries and most airlines (Kanki et al., 2010). Given that airlines are often regarded as having exemplary safety records, it is not surprising that other industries have started to adopt CRM training as well (Salas, Wilson, Burke, Wightman, et al., 2006).

However, evaluative research on different CRM programs has not often found CRM programs successful. Reviews on the subject have not found a clear link between CRM and a reduction of accidents (Salas, Wilson, Burke, & Wightman, 2006). A key problem is that evaluative studies on CRM suffer from a variety of methodological issues (Edkins, 2002), including selection of outcome measures. When other industries adopt CRM, this makes it hard to recognise successful CRM programs, and to understand which parts of a CRM program are important for success, or even to interpret what "success" looks like outside of encouraging specific cockpit behaviours.

This leads to our main research question:

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

To understand the adoption of CRM in an industry, it is first important to understand what the adopters understand CRM to be, as different people can conceptualise CRM training differently. Changes in conceptualisation or definition during adoption of CRM should then be reflected in the evaluation of CRM programs (Holt, Boehm-Davis, & Beaubien, 2001). Therefore, we have split the main research question into:

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

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

2.5 Methods

The purpose of this systematic literature review was to explore how industries have adopted CRM. Where applicable, we followed the guidelines of the PRISMA statement (Moher, Liberati, Tetzlaff, Altman, & The PRISMA GROUP, 2009) on the reporting of systematic reviews. Our goal was not to assess the efficacy of CRM as a thing-in-itself, rather it was to study how CRM has been adopted and evaluated, therefore the elements of the PRISMA guidelines that are focussed on reaching conclusions about efficacy did not apply. In addition, to lower the cognitive burden on the reader, we describe the review of the conceptualisation and evaluation sequentially, and draw conclusions after each section, before combining the results.

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

2.5.1 Search process

To orientate, we familiarised ourselves with the CRM reviews and textbooks on CRM (Kanki et al., 2010; O'Connor, Campbell, et al., 2008; O'Connor, Flin, & Flechter, 2002; Salas, Wilson, Burke, & Wightman, 2006). From the orientation, we established the initial search terms and explored the domains that were available. We chose to analyse the following fields: the maritime industry (MAR), the nuclear power industry (NPI), the oil & gas industry (O&G) and air traffic control (ATC). Industries such as firefighting and railway were left out because of the low number of publications. We also excluded military CRM and healthcare CRM because these industries have already been thoroughly and recently reviewed (Buljac-Samardzic, Dekker-van Doorn, van Wijngaarden, & van Wijk, 2010; Kanki et al., 2010; O'Connor, Hahn, & Salas, 2010; O'Dea, O'Connor, & Keogh, 2014; Salas, Bowers, & Cannon-Bowers, 1995; Weaver et al., 2010).

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

- MAR: ("crew resource management" & maritime) OR ("bridge resource management") OR ("bridge team management") OR ("maritime resource management") OR ("engine room resource management")
- NPI: ("crew resource management" & "nuclear power")
- O&G: ("crew resource management" & oil & gas) OR ("crew resource management" & offshore) OR ("deep water well control") OR ("crew resource management" & "well operations")
- ATC: ("crew resource management" air traffic control) OR ("team resource management") OR ("controller resource management") OR ("air traffic team enhancement") OR ("controller awareness and resource training") OR (ATCRM)

2.5.2 Selection process

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

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

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

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

2.5.3 Conceptualisation review

Analysis

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

Goals

First, we looked at the major goals of CRM training. For this we went to Lauber's (1984, p.20) definition of CRM as "using all available resources — information, equipment, and people — to achieve safe and efficient flight operations". This is the oldest definition of CRM and variations of this definition continue to be used. The "using all available resources" component is open to multiple interpretations, however "safe" and "efficient" provided two clear starting categories. We analysed all papers based on whether they mentioned safety and efficiency as goals. Safety included reduction of accidents and incidents. As efficiency might not be easily applied to all industries, we translated this to any increase in performance. Statements regarding desired, expected, and found effects were all counted as indicators of goals.

Levers of change

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

- the unit of analysis for the immediate effects of CRM (Flin, O'Connor, & Crichton, 2008) and
- the stance of CRM with respect to compliance

Flin et al. (2008) highlighted the unit of analysis (e.g. individual or team) as a contested issue in the conceptualisation of CRM. Changes in attitudes, knowledge, or skills of the trainees move with the individual as they are put in different situations and teams, suggesting that CRM operates at the level of individuals. Familiarity built among team members, however, would disappear if a team does not stay together, therefore some benefits of CRM may be considered to reside at the team level.

We first listed each of the identified effects in each paper, and categorised these according to 'individual', 'team' or 'other'. Effects that could not be placed in the individual or team category were categorised into the 'other' level, as were changes to procedures, the work environment, or 'culture'.

Across the CRM literature, multiple stances regarding compliance can be found. Helmreich, Klinect, and Wilhelm (1999) see both intentional and unintentional incompliance with procedures as being both a potential cause for accidents, and something that CRM training could help avoid and manage. On the other hand, Flin et al. (2003) stated that CRM is about non-technical skills, which are not directly related to standard operating procedures. Haller and Stoelwinder (2013) went even further and argued that CRM is an alternative to a focus on operating procedure compliance. In most cases papers do not explicitly state their view on compliance. Because of that, we reduced the analysis to three categories: whether papers mentioned compliance should increase (or violations decrease), whether papers mentioned that CRM is not related to standard operating procedures, or whether there was no explicit mention of compliance to standard operating procedures.

Scope of CRM

Finally, we considered the scope of CRM training. CRM training in the cockpit does not always focus on the same elements. For example, Salas et al. (1999) mentioned that CRM is about coordination, and stated that CRM is about improving team competencies. Helmreich, Klinect, and Whilhelm (1999) on the other hand stated that improving teamwork is important, however, is not the primary goal of CRM. To assess the scope of CRM programs in different industries we analysed papers based on their claims of what type of errors CRM helps manage, and the modules included in the program.

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

To analyse modules, we considered only papers that reported the module names in a CRM program, and combined different papers that reported the same CRM program. As there is no standardisation in modules names, we categorised the modules into the NOTECHS (Avermaete & Kruijsen, 1998) skills to allow comparison. We added ‘personal resources’ and an ‘other’ category to the NOTECHS list (Table 2). “Personal resources” was included because this was mentioned in the original NOTECHS report, but excluded from the final document because it was considered too hard to observe for a behavioural marker list. The “other” category was added, because some training modules might not fit within the NOTECHS classification. If a module name fitted within two NOTECHS skills, it was counted as half a point for both skills.

Table 2. NOTECHS classification for CRM modules

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

Synthesis

For the synthesis of conceptualisation data items, we combined the single paper analyses into counts per industry and percentage of industry total. For CRM modules, in the calculation of percentage per industry, the total per industry only papers that specified training modules are counted. In addition, since multiple modules can cover the same skill, we calculated how many CRM training programs covered a skill at all, and how much of the training was dedicated to each skill per industry on average. To calculate how much training was dedicated to a skill, we calculated what percentage of modules within each training program fell within each NOTECHS skill, and averaged this out for all studies that specified modules per industry.

Bias

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

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

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

2.5.4 Evaluation review

Analysis

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

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

Synthesis

For the synthesis of the evaluation data, we ordered the claims and evidence into a grid that was sorted by industry, evaluated variable, and direction of the results. We categorised the effects of CRM training using the Kirkpatrick levels of training evaluation (Kirkpatrick, 2016). However, as we included comparison as one of the requirements for evidence, we left out the reaction level, as reactions are not evaluated in comparison to anything. Theoretically it would be possible to go beyond the Kirkpatrick levels, and look at industry-wide changes in accident causation after the introduction of CRM. However, the subjectivity involved in accident cause determination (Lundberg, Rollenhagen, & Hollnagel, 2009; Rasmussen, 1990; Wrigstad, Bergström, & Gustafson, 2017) makes it impossible to meaningfully review this level and is beyond a review on how CRM has been adopted.

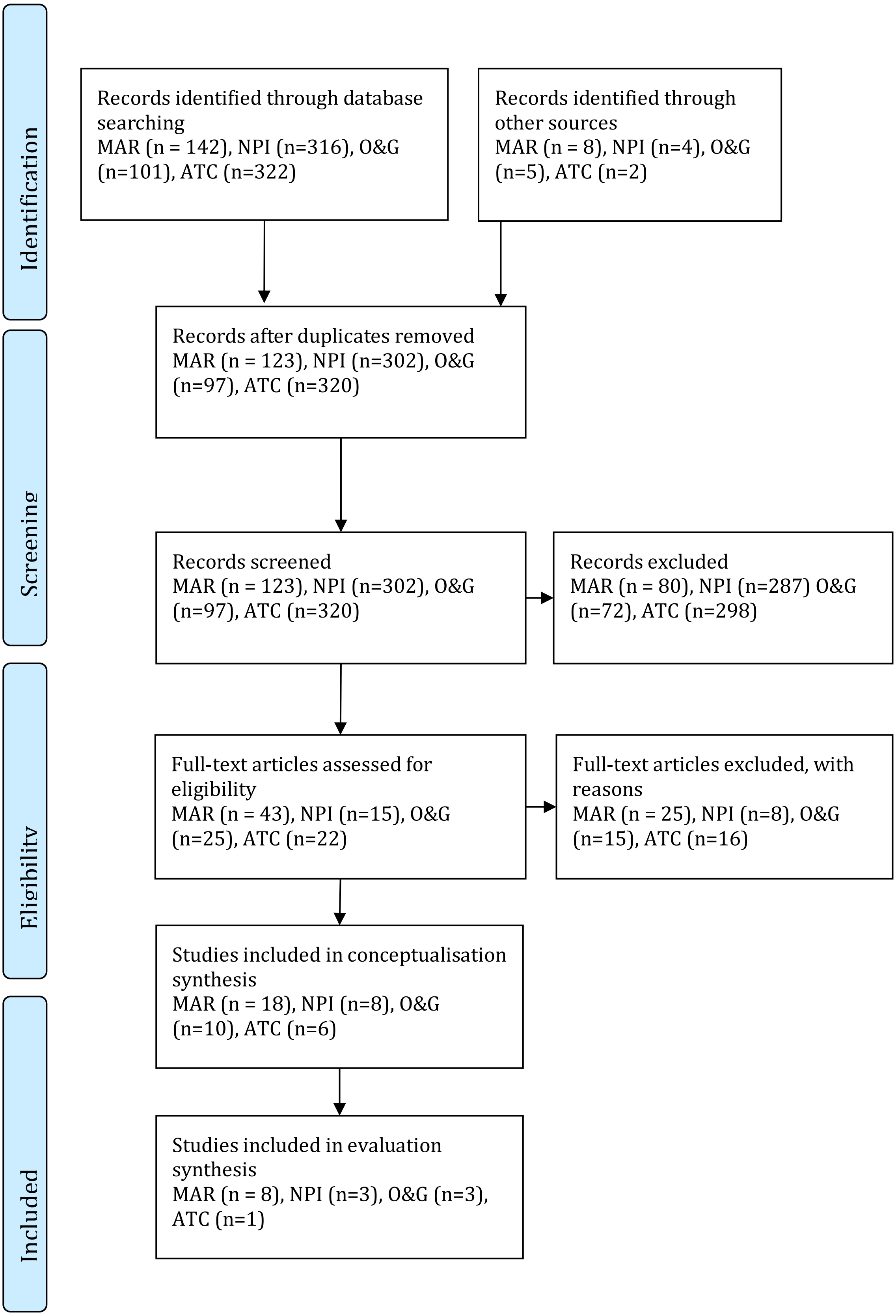
Bias

Our evaluation could be biased towards showing CRM as being more often effective than it usually is. It is likely that only the more thoroughly developed the CRM programs are evaluated, which filters out many less thorough and less well-funded CRM programs. In addition, evaluations that show CRM as effective are more likely to be published than studies that did not find CRM to be effective. This means we expect to review the higher quality CRM programs in the reviewed industries. This does not affect our research question directly, as we are interested in how CRM has been adopted rather than its efficacy per se; but it does mean our summary of how CRM has been evaluated may misrepresent how effective CRM has been on average.

2.6 Results

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

Figure 1 PRISMA flow diagram



2.6.1 Analysis and synthesis for conceptualisation of CRM

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

Organisational goals

Lauber (1984) stated that the goals of CRM were to achieve safe and efficient operations. We analysed whether the goals 'safety' and 'efficiency' were returned in all analysed industries. Table 3 shows the percentage of papers per industry that mention either of these goals in relation to CRM training. We interpreted safety as an avoidance of negative events (accidents and incidents) and translated efficiency into more general performance improvement, as efficiency might be hard to apply to certain industries.

Table 3. Percentage (number) of papers that mention the specified organisational goal per industry

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

The papers linked CRM training to the safety goals and to the performance goals of an organisation in all industries. All but one paper mentioned improving safety as a goal (98%)[1]. Most of the papers also made the link between CRM and improving performance (81%). There are, however, some differences between industries in this regard: in the NPI and ATC industries all papers mentioned performance goals, in contrast to the MAR (78%) and O&G (60%) sectors. Although this difference in percentages might reflect a difference in emphasis of CRM objectives between industries, it can also be caused by the sample, as the papers not mentioning performance tended to be short papers.

[1] The exception is a paper from maritime investigating the relation between shared mental models, CRM and performance (Brun et al., 2005), arguably making safety not the topic of the paper, but still a goal of the training program.

Levers of change

Not all CRM programs are designed on the same philosophy and do not expect to reach their goals in the same way. For example, Salas et al. (1999) mentioned that some programs focussed on attitudes and others focussed on knowledge or skills. Salas et al. (1999) also commented that this difference was often hard to distil from papers. To assess differences in the envisioned levers by which CRM affects the organisational goals, we analysed papers for indications of their unit of analysis, and whether increased compliance was a mechanism to achieve the larger goals.

Unit of analysis

All papers were analysed on the units of analysis behind desired or claimed effects of CRM training. Table 4 shows what percentage of papers mention effects on the individual level, the team level, and other levels of change.

Table 4. Percentage (number) of papers that link CRM to effects at individual and team level

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

The analysed papers mentioned team effects much less frequently than individual effects (14% versus 100%) in relation to CRM training across all industries (0% - 20% for team effects). Examples of team level effects of CRM include team bonding (Bailey, 2014), team-member specific knowledge (Smith-Jentsch, Baker, Salas, & Cannon-Bowers, 2001), and shared mental models (Brun et al., 2005).

In terms of other changes linked to CRM, we found bigger differences between industries. In MAR (6%) few "other" effects were mentioned, while in O&G (60%) most of the papers mentioned effects that did not directly reside in individuals or teams. Examples of changes at other levels include culture (Davis, 2001; Mearns, Whitaker, Flin, Gordon, & O'Connor, 2003), procedure changes (Grinde, 1994), (practical) changes to the work environment (Davis, 2001; S. K. Kim & Byun, 2011), aligning teams with each other (Bailey, 2014), testing out plans (Crichton, 2009) and discovering training requirements (Wang et al., 2014). These do not all have the same unit of effect. Changes in culture or procedures are expected to affect new workers, even if they have not had CRM training. The benefits from aligning teams only remain if the same teams continue interacting with each other.

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

Compliance

The papers were analysed on the question of whether CRM achieves its goals because it leads to better procedural compliance. Table 5 documents the percentages of papers per industry that link CRM to compliance with procedures, or that state CRM is not about procedures.

Table 5. Percentage (number) of papers that link CRM to compliance or that state CRM is beyond procedures

Less than a quarter of the papers (21%) explicitly mentioned that CRM training was not about operating procedures. There was some variation between industries (0% - 38%), with ATC scoring the lowest. An approximately equal amount of papers linked CRM to adherence to procedures (24%). Within an industry, the difference was never bigger than one paper. However, most papers did not discuss whether there is a relationship between CRM and procedures (55%), especially in O&G and ATC (80% - 83%). Two papers, one from O&G (Grinde, 1994) and one from MAR (Wang et al., 2014), mentioned that CRM training was used to change procedures, but these seem to be exceptions. Overall, there are differences between CRM programs, but there were no tendencies between industries.

Scope of CRM

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

Error specification

For flight crews CRM was introduced in response to errors that were observed in accidents. However, this approach does not have to be universal, nor do programs necessarily focus on all error types. Table 6 shows the percentages of papers per industry that mention reduction of error as a goal in relation to CRM training. Furthermore, the table shows the specification of types of errors CRM addresses, as an indication as to where the emphasis is in different sectors.

Table 6. Percentage (number) of papers that mention reduce errors per industry

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

Generally, the papers linked CRM training to a reduction in errors (83%). In ATC (100%) all papers mentioned errors in relation to CRM; in the other industries this percentage was slightly lower (75%-83%). In addition, as expected, technical errors were mentioned in only a few papers on CRM: between 0 and 30% with an average of 14%. In contrast, a reduction in both human and team errors was associated with CRM training in half or more of the papers in each industry; except for team errors in MAR (16%). CRM was strongly related to a reduction in team errors in the NPI and ATC industries (63% and 67% respectively, in each case larger than the score on human error of 50%). Note that there is some ambiguity in classification: some authors saw team errors as a subset of human errors (Mearns et al., 2003), while others suggested they are separate categories (S. K. Kim & Byun, 2011). These differences in categorisation made it hard to draw conclusions; however, it seems NPI & ATC

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

put most emphasis on the team aspects, and in O&G the team aspect was only one of the addressed aspects.

Course modules

To look at the scope of CRM programs we searched for the modules included in training programs. To be able to compare industries, we categorised all modules along the NOTECHS skill list (Avermaete & Kruijsen, 1998).

Table 7 shows the percentages of programs that include a module that covers (part of) a NOTECHS skill, as well as what percentage of the modules within a program are about each skill on average per industry.

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

Table 7. Modules in CRM training by NOTECHS categorisation

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

The variation in CRM programs resembled the conflict in the theoretical literature. The skills 'decision making' (60%) and 'situation awareness' (50%), were reflected only in about half of all programs. All of the analysed programs in

NPI and ATC had modules on these skills, however, both only have two papers that described the modules in the program.

'Leadership & management' (95%) and 'cooperation '(95%), the skills most clearly related to the team element, were covered in practically all CRM training programs. Only one CRM training program from O&G did not have modules that would fall into these two skills. When looking at how much of each training was dedicated to each skill, it becomes even clearer that these skills received the most emphasis, as programs often had multiple modules that fell within these two skills. In most industries, more modules covered cooperation elements than leadership & management, with the exception of MAR. This could reflect a more clearly defined hierarchy and role distribution in MAR than in other fields.

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

More than half of all papers (65%) included modules that did not fit into the NOTECHS 'plus personal resource' categorisation. These non-fitting modules were least common in O&G (20%), and most common in MAR (82%) and NPI (100%). When looking at the modules in the "other" category, most were unique cases. Topics that returned multiple times were; emergency management related modules in four MAR (36%) trainings; error and accident analysis in three MAR (27%) trainings and in one NPI training; and Attitudes in two MAR (18%) trainings and one ATC (50%) training. The error and attitude modules were likely framing modules, rather than representing different training objectives. The fact that emergency management often received its own module in MAR, suggests that there people see emergency interactions as fundamentally different; which is supported by a comment that hierarchical leadership works for normal situations on ships, but a flatter hierarchy is required in emergencies (Habberley, Barnett, Gatfield, Musselwhite, & McNeil, 2001).

These results show that there is a general agreement that the team element is part of CRM. It is likely the case that many training programs see CRM as more than just about the team element, but the size of this group cannot be drawn from our analysis. Certain programs discussed fatigue and self-management (Byrdorf, 1998), while other programs mentioned how stress can be recognised in other people (Smith-Jentsch et al., 2001), relating stress to the team element. Similar to personal resources, decision making and situation awareness can be addressed for individuals or for teams. Thus, the inclusion of personal resources, decision making, or situation awareness does not necessarily mean CRM training aims to target non-team processes.

2.6.2 Conceptualisation synthesis summary

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

2.6.3 Analysis and synthesis for CRM evaluation

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

Maritime

Learning

In MAR, five studies examined learning effects. Fonne and Frederiksen (1995) found that attitudes changed directly after the course, and moderated over a period of six months. However, they only reported percentages and did not do any statistical tests to rule out random variation.

Inoue and Takahashi (2005) used one item attitude-like measures, of which they compared the averages, and found that junior trainees changed their attitude more than senior officers did. They did no statistical test on the before/after changes, or between groups, to rule out random variation.

Saatcioglu, Asyali, and Cerit (2005) claimed that trainees changed their attitudes towards cooperative learning. However, they tested for statistical difference per item of the questionnaire, not per the attitudes that make up cooperative learning.

Håvold, Nistad, Skiri, and Ödegård (2015) investigated how CRM training achieved its effect. Using a questionnaire, they found that people who thought the course was better set up, also believed their knowledge, skills and attitudes changed more. In addition, those who believed their knowledge, skill and attitudes had changed more, also thought their behaviour had changed more. However, the questionnaire was only assessed for reliability, but not validity, and it was a multiple-regression study, not a comparison study.

Brun et al. (2005) looked at the effects of CRM training on shared mental models. They found no difference between the experimental and control group. The teams in the study had already worked together before the training, which can explain why no change was found. Cockpit CRM has usually been tested with newly formed teams. The test was too small to expect any effect of statistical significance, therefore, the results give no indication whether CRM has an effect on shared mental models in already established teams.

None of these studies met our evaluation criteria.

Behaviour

Four papers looked at how CRM training was applied by the trainees and led to behavioural changes. Byrdorf (1998) and Håvold et al. (2015) used questionnaires and interviews to ask people whether they had changed their behaviour afterwards. In both studies, people reported changed behaviour, but neither study did a comparison, either a 'before-and-after' or with a control group. Brun et al. (2005) and Wu, Miwa, Shimamoto and Uchida (2015) both compared two teams per condition and found small differences. Brun et al. (2005) doubted whether the differences were caused by CRM, while Wu et al. (2015) explained this as a small effect size from CRM training. Both studies were too small for a statistical test. None of the studies met our evaluation criteria.

Organisational effects

In terms of organisational goals, two MAR studies reported effects. Byrdorf (1998) found big changes on multiple organisational measures when comparing 1992, the year the intervention started, with 1996, four years after they started doing the training. They found large improvements on multiple measures. However, there was no control group and only 2 years were reported, taking away the opportunity to rule out alternative explanations. Other explanations have some merit, as there were three other safety interventions being implemented in the company (Byrdorf, 1998), and there was a general trend of reduction of accidents in the industry at that time (Habberley et al., 2001).

Wang et al. (2014) claimed success of their CRM training because of changes they made to their planned operations. This included changes to procedures, interface design, safety margins, the flow of information and identified additional training needs. These outcomes make CRM more into a test of plans, as opposed to the development of skills. These outcomes can be as useful for both safety and efficiency, but it is impossible to verify whether these suggestions improved safety and efficiency.

Neither study met our evaluation criteria.

Nuclear power industry

Learning

Two studies from the NPI looked at learning effects of CRM training. Harrington and Kello (1992) used a questionnaire based on the Controller Resource Management Attitude Questionnaire (CMAQ) and found an effect on all of its attitudes, recognition of stressor effects, communication & coordination, and command responsibility.

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

The results of Harrington and Kello (1992) met our evaluation criteria and support attitude change from CRM, but the results of S. K. Kim and Byun (2011) on attitudes did not meet our evaluation criteria.

Behaviour

S. K. Kim and Byun (2011) compared a CRM trained group and control group on multiple measures in a simulator. They measured situation awareness and workload through self-report, for which they found no effects. For team behaviour, they used behavioural markers using evaluators and found a significant difference there between groups. Their conclusion was that CRM mostly works at the team level, not on individual skills. However, the groups were not randomly selected; the CRM trained group was more experienced and familiar with operating a different type of nuclear reactor. In addition, the evaluators were also the designers and trainers of the course, which could have biased the ratings. This evaluation met our evaluation criteria, although alternative explanations for the effects do exist.

Organisational effects

Davis (2001) found that a CRM program led to improvement in operational systems, attitudes, culture, human performance and reportable events associated with human performance and supervision. However, the numbers were not reported, nor did the author how they were measured. With that, the study does not meet our evaluation criteria.

Oil & gas

Learning

One study in O&G evaluated learning effects. O'Connor and Flin (2003) looked at the learning in terms of attitudes and knowledge of the trainees before and after the training. For attitudes they used a Controller Resource Management Questionnaire-based questionnaire, which they reduced from the original six factors of the questionnaire to four. Of these four factors, decision-making and personal limitations changed in the desired direction after the training; but there was no change for situation awareness and communication. The researchers reported multiple issues with their test, which can provide explanations for the lack of effect of the training; 1) because of the high initial scores, 2) the small sample size compared to the expected effect size, and 3) the reliability of the scales was not great. In their knowledge test they found no changes. A possible reason given for this was that practitioners might not have been that motivated for the last test, as they were about to go home. The study met our selection criteria and provided mixed support for CRM training changing attitudes.

Behaviour

The only study that evaluated CRM training on the level of application comes from Moffat and Crichton (2015). They looked at behavioural change from teams during multiple exercises in CRM training. They observed changed scores between the earlier and final exercise. However, the data was limited to two exercises, and the marker system was still in development, meaning the scale had not been tested on reliability or validity. It is unclear whether the evaluators were also the trainers of the course. The study can be expanded, however, in its current form the study did not meet our selection criteria.

Organisational effects

For organisational effects, Crichton (2009) mentioned that a trained team had an exemplary performance in terms of non-productive time, health and safety data and project goals. However, the measures were not reported and the sample size was one. The study sounded promising, but the study did not meet our evaluation requirements.

Air Traffic Control

Learning

There was only one paper that evaluated CRM training in ATC. It was a larger EUROCONTROL project that tested a program that was rolled out in multiple European countries (Woldring & Isaac, 1999). This study used a Flight Management Attidue Questionaire-based questionnaire to test for attitudes before and after the training. For all items combined they found a statistically significant change. On another smaller sample they found statistically significant change on 17 of the 38 items, however, some of these items changed in the undesired direction. The study lacks a comparison per attitude, which the items are supposed to reflect. Because of that, the study does not meet our selection criteria.

2.6.4 Evaluation synthesis summary:

A summary of the evaluation in different industries can be found in Table 8. There were six evaluations that met our quality criteria, of which two were positive, one was mixed, and three showed no effect. This means in the most rigorous tests, CRM was effective less than half of the time. Both tests with positive results came from NPI, however, NPI also had two tests showing no effect. In O&G there was one test with mixed results and one test that showed no effect. In MAR and ATC none of the analysed studies met our criteria. In all of the cases that met our evaluation criteria, the measurements were done right after the training. This leaves the question as to any of the few found changes are lasting.

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

Table 8. Summary of evaluation claims and results

		MAR			**NPI**			**O&G**			**ATC**		
		pos	mix	no	pos	mix	no	pos	mix	no	pos	mix	no
Attitudes	Tested	-			1				1		-		
	Claims	4			3				1		1		
Knowledge	Tested	-								1			
	Claims	1								1			
Skills	Tested	-											
	Claims	1											
Shared mental models	Tested			-									
	Claims			1									
Situation Awareness	Tested						1						
	Claims						1						
Workload	Tested						1						
	Claims						1						
Behaviour	Tested	-	-		1			-					
	Claims	3	1		1			1					
Organisational	Tested	-			-			-					
	Claims	2			1			1					

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

2.7 Discussion

2.7.1 Summary of evidence

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

2.7.2 Limitations

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

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

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

2.7.3 Conclusions

Previous reviews of CRM have argued for more evaluation, especially across multiple levels (individual, team, and organisational effects). We agree that showing a decrease in accidents is one useful outcome measure, but this is probably not enough. Evaluation needs to focus on the process through which CRM programs lead to change. This is not something that can be assumed from the label 'CRM', because of the differences between programs in how CRM is conceptualised. Each CRM study needs to make their hypotheses about how CRM creates change explicit. Salas, Wilson, Burke, and Wightman (2006) argued for standardisation in training and evaluation of CRM. That might have made sense once, but considering CRM's long history and diverse fields of application, we now argue for the opposite. We suggest researchers go behind the label CRM and describe what they intend to do and expect to change. This can make the differences that are normally hidden behind the label CRM visible and create a more auspicious way to evaluate the effectiveness of CRM in all its diversity.

For example, there is good indication that the training model that is reflected in the Kirkpatrick evaluation levels does not work well for CRM. This model suggests that training leads to learning with changes in knowledge, skills, and attitudes, which leads to a change in behaviour, leading to improved safety and performance. Of all these elements, it is most supported by research that CRM leads to changes in attitudes and behaviour, but the links between these elements are not well established. Two studies have touched on the link between behaviour and attitudes, but do not provide strong evidence for this link. The oldest paper was about pilots (Helmreich, Foushee, Benson, & Russini, 1986) and has methodological shortcomings. First, it looked at separate items of a questionnaire, not the attitudes they represented. Second, the predictive values were based on a test with the same data set the model was built with, instead of testing it on a new data set. A more recent paper from MAR (Röttger, Vetter, & Kowalski, 2013) found no link between attitudes and behaviour, but further exploration of the data showed there might be a relation between low-mid attitudes to behaviour. As far as there was a link between attitudes and behaviour in that study, it was a universal link of all attitudes to all behaviour, not a link of specific attitudes to related behaviours. This can suggest that any link may be through a single overarching variable, such as psychological safety climate (Bradley, Postlethwaite, Klotz, Hamdani, & Brown, 2012), and not by a change in attitudes.

This leaves two possibilities. The first is that there is some link between behaviour and attitudes, but it is a weak one. Considering that in all reviewed studies, the biggest effect on a single attitude found was 9%, which is less than 0.5 points on a 5-point Likert scale. When that is combined with a weak link to behaviour, it is unlikely CRM meaningfully changes behaviour. It is possible that the claims about the success of CRM do reflect real change, but that it is not captured by this model and that it comes about through other processes, such as change in the psychological safety climate. In either case, it means that CRM training programs focus on elements that have weak effects, and that training could be optimised by spending time on different elements. To find answers to these questions, however, CRM researchers need to explicitly describe how they expect change to take place after the training and test accordingly.

2.8 Outcome for the research aims

The review found CRM programs diverse in conceptualisation, both within industries and between industries, and without substantial evidence that CRM was making a difference. Because of the diversity in the conceptualisation of CRM, there was also no clear direction as to what should be investigated to answer the question of whether CRM is effective, let alone what kind of CRM training could be effective for blast crews. These results showed that there was no agreed upon state-of-the-art way of organising teams and that there was no clear evidence that teaching teams how to organise is effective. This pointed to a gap in the literature on how teams organise, how an organisation comes about, and what influences how teams organise.

Chapter 3: **Studying everyday work for safety purposes**

3.1 Statement of contribution to co-authored published paper

This chapter includes a co-authored paper. The copyright has been transferred to Taylor and Francis, however, authors have the right to include their article in a thesis or dissertation. The bibliographic details of the co-authored paper, including all authors, are:

Havinga, J., Dekker, S., Rae, A. (2017). Everyday work investigations for safety. *Theoretical Issues in Ergonomics Science*, (July). http://doi.org/10.1080/1463922X.2017.1356394

My contribution to the paper involved:

The collection of fieldwork data and its analysis; The conceptualisation of the problem, paper's goals, and questions with normal PhD supervision from Dr Rae; The review of the literature; The first creation of the first draft; The editing of the draft based on feedback from Prof Dekker and Dr Rae.

(Signed) ______________________________ (Date)____________

Joseph Paulus Havinga (Corresponding author)

(Countersigned) ________________________ (Date)____________

Dr Andrew Rae (Principal supervisor and co-author)

(Countersigned) ________________________ (Date)____________

Prof Sidney Dekker (Principal supervisor and co-author)

Observations used in this study were conducted under ethics registration GU Ref No: 2015/191

3.2 Rationale for the study

There was no sound basis to learn from state-of-the-art ways teams organise, therefore more emphasis fell on learning how blast crews organise their work. My initial observations of blast crews further confirmed that the majority of Crew Resource Management (CRM) literature was unsuited for blast crews, as the way blast crews coordinate their work was different from any description found in the CRM literature. CRM courses often include modules on how to communicate with team members, suggesting how to raise issues, how to question team members, or how to make decisions collectively, and emphasise the need to build shared mental models. My initial observations with blast crews, however, showed a different picture. During the job, practitioners would hardly talk to each other about the job. Often practitioners could not even talk to crew members as they were too far away from each other. Nevertheless, within the blast crews everyone appeared to know what to do, adapted to disturbances, and helped fellow team members out. A lot was going on during this everyday work that did not resemble anything I found in my search through the CRM literature. This combination of a (seemingly) smooth and efficient work process, a process that was different from what I encountered in the literature, and a safe process according to the accident rates, strongly indicated that there was something that could be learned from how blast crews organise their work.

Investigating to learn how work happens, rather than focussing on accidents, fits in with both HRO and Resilience Engineering. Arguably, this shift towards normal work had began earlier with the work of Turner in 1978 saw accidents as by-products of normal organisational processes in his man-made disaster theory (Woods, Dekker, Cook, Johannesen, & Sarter, 2010). However, despite that this idea has been around for some time, there are no clear instructions on how to design a study of everyday work and make it relevant for safety, or just relate it to safety. For this, I reflected on the possibilities and consequences of these choices for my study, using my initial observations of blast crews.

Everyday work Investigations for Safety

3.3 Abstract

A key theme in safety is the investigation and improvement of work processes and conditions. This is built on a tradition of studying accidents, but there is growing interest in the investigation of everyday work. Researchers have started using terms such as success, normal, and everyday work investigations, but behind these can be different activities. This paper makes explicit some considerations behind investigations of everyday work, how these considerations shape an investigation, and how they reflect different schools of thought. These considerations can be used as an aid for designing everyday work investigations for safety and serve as a first step towards critical reflection on investigating everyday work.

3.4 Introduction

The safety community is seeing a growing interest in investigating success and normal work, in addition to (or instead of) investigating things that have gone wrong, such as incidents and accidents. What has been called "Safety-II" specifically advocates looking at how success is created, instead of how a particular failure has happened (Hollnagel, 2014). In the study of work itself, this is nothing new of course. Taylor's and Gilbreth's time and motion studies investigated 'normal work', though their concern was productivity, not necessarily safety (even though some of their studies did include a focus on the errors made by the worker) (Taylor, 1911). Ergonomics, Naturalistic Decision Making (NDM), distributed cognition, cognitive work analysis and Cognitive Systems Engineering (CSE) have since all contributed to the study of everyday work (Hoffman & Militello, 2008). In addition, there are many forms of task analysis, some of which only focus on how work should be done, rather than on how it is actually done (Annett, 2000), and safety is not always, or is seldom, the main concern. The challenge is in talking about everyday work and safety at the same time.

From a safety perspective, the importance of understanding everyday work relates to the mismatch between how people think work is done—how it is captured in procedures—and how work is actually done (Hollnagel, 2014). This has been called the difference between task and activity (Leplat & Hoc, 1983), procedure and practice (Dekker, 2003), or indeed work-as-imagined and work-as-done (Dekker, 2005a; Hollnagel, 2014). In more judgmental terms, actual work that does not match work as imagined has also been characterised as a workaround or even a violation (Debono et al., 2013; Reason, 1997; Spear & Schmidhofer, 2005). Where cognitive work takes place 'in the wild,' as Hutchins (1995) puts it, pragmatic and highly nuanced adaptations of tools, techniques, and practices is not only extremely common, but generally crucial for the creation of safety at the sharp end. Hollnagel (2012) too, stresses that work-as-imagined can never capture the full complexity, or the 'messy details' (Nemeth, Cook, & Woods, 2004) of actual work. This means that addressing the mismatch is not a matter of enforcing closer alignment between actual activities and how work was once imagined to be. Rather, the challenge for an organisation becomes to learn more about how it actually operates mostly successfully and safely at its many "sharp ends" of practice (Woods et al., 2010). For the purposes of this paper, everyday work investigations refer to organised activities to learn how people or organisations operate. This is done to better understand how safety is created and broken in people's everyday activities, particularly in how they respond to goal interactions, local insights and resource limitations.

The historical legacy of safety research and practice has created methods, guides and critical reflections on investigating things that have gone wrong—many more than there are for investigating everyday work. Even though accident and incident investigation techniques might suggest some basis for investigating everyday work, they provide little practical guidance on how to do so. Ergonomic based methods, on the other hand, often come from a background of looking at individuals and an aim for the redesign of technology. While useful, these are not the only relevant factors for safety. One can set out and apply these investigation methods and tweak them until they produce results in the new setting. However, this approach would miss out on more fundamental considerations of the worldviews, goals, and concepts that come with a focus on accidents. This could limit the investigation of everyday work, and be unable to cope with the challenges the come with investigating everyday work. For one, many work-as-done adaptations are integrated into practice so smoothly, subtly and unobtrusively that they are all but invisible to the untrained eye (Cook & Woods, 1994). Without the rubble of an accident or incident to start backtracking from, there is no clear trigger point for investigating everyday work.

In addition, everyday work investigations are beset by many of the same problems as incident/accident investigation. The cognitive systems in which everyday work is embedded are complex, deeply intertwined and filled with uncertainties, and investigator observations are inevitably driven by the categorical forestructures that are taken into the field of observation. That said, there are distinct advantages of investigating normal work. There are more opportunities to gather data (Hollnagel, 2014) and data is not limited to retrospective records and interpretations of an event. Observations can be trained on ongoing work instead. Opportunities thus exist to both generate and verify hypotheses with new data. Of course, without a bad outcome that triggered it, the investigation of normal work should be free from the psychological defense mechanisms (and from at least some of the effects of outcome and hindsight bias) that typically mar the aftermath of an incident or accident (Dekker, 2014b; Merry & McCall Smith, 2001).

Our purpose in this paper is to discuss the theoretical considerations for conducting everyday work investigations related to safety. In particular, we discuss the effect that different research questions and methodological assumptions have on the type of conclusions that can be reached, and the relationships between these conclusions and the ultimate goal of improving workplace safety.

To illustrate our discussion, we use as a running example an empirical study of everyday work by blast crews. It is not our intention to provide a full report of the methods or results of that study here, but to use our experiences in designing and conducting the study to support a broader reflection on what can be investigated and what makes sense to investigate. We also reference other everyday work studies, adopting a range of different approaches. Using these examples, we discuss some considerations to make when investigating everyday work, linked to various safety schools of thought and fields that study work.

3.5 Blast crews

In a mine, the role of the blast crews is to ‘soften’ up the ground. By using explosives, they crush the ground, so that diggers in the mine have an easier job. If their blast is too weak, the rocks afterwards are too big. If their blast is too strong, environmental norms can be exceeded and debris can be flung away from the blast area. Besides debris, the job of blast crews is safety critical due to risk of premature and uncontrolled detonation, and managing risks that come with operating heavy vehicles.

Before a blast can be done, the area to be blasted is surveyed. After it is surveyed a ‘shot’ is designed, which means a plan is made to decide what kinds of explosive ‘product’ will be used and how the product will be distributed within the area. According to the design, the holes are drilled where the product will be loaded into. The next part is the loading phase where the holes are filled. Here the crew first ‘baits’ the holes by lowering a booster — a small cylindrical explosive — attached to a detonation cord into the hole, then the product is loaded into the hole and the hole is topped off with stemming. The booster is attached to a detonation cord that sticks out of the top of the hole. The last step is to connect the cords according to the plan, before the shot is set off from a safe distance.

3.6 Our investigation

Our study was part of a larger a problem-orientated ethnography. However, in this paper, we only report on the orientation part of this study, as we use it as an example to reflect on how everyday work can be studied. For this, the first author visited blast crews and followed them around for multiple workdays in a row. The researcher had received basic training and done the necessary inductions to work in a blast crew, but had no prior experience in the field. If the management of the quarry or mine gave the researcher authorisation to work, he would help the crews out during work. Otherwise, he would only observe, ask questions, and take notes. During the work, the crews were observed and short questions were asked. During breaks, the researcher did formal interviews with the crew members.

The observed blast crews varied in size between three and ten people. There was always at least one shotfirer that was in charge of the crew. Each crew had at least one truck operator and one bench hand that did extra work on the ground. During preparations and the baiting of the holes, there is only a loose team structure with few pre-delegated tasks. The holes during the observed work varied between 3 and 15 meters deep and the number of holes per shot varied from 50 to 350.

An aim of our study was to investigate the possibilities of Crew Resource Management training for blast crews in mining. Because of this, we were mostly interested in teamwork processes and the interactions between crew members and agents outside the team. During this fieldwork, the researcher observed surprisingly little explicit task-related verbal communication. Yet, the crews seemed to be working efficiently and showed other teamwork processes such as mutual performance monitoring and back up behaviour (Salas, 2005). Upon closer inspection, the researcher found that practitioners purposely left 'traces' for themselves and other crewmembers in the environment. A lot of communication was happening nonverbally, although it was not necessarily implicit.

Some of the practices were specific to teams, others were more universal. We do not know whether these practices have evolved naturally or whether they had been designed (Hutchins, 1991), but they did not come from the company's rules, procedures or formal training. These interactions were part of the process of how work gets done, but were not part of work-as-imagined by the blunt end, and we think they are unlikely to be found in an accident investigation. Therefore we see it as a good example of what everyday work investigations can find and how it will lead to different questions than accident investigations. We have chosen three of these interactions observed in one crew as examples to use in further reflection on how to shape further investigation.

3.6.1 Observed blast crew interactions

Hose placement as a beacon

The truck with the product needs to move many times for the hose to reach all the holes. The number of times the truck is moved is kept as low as possible. This reduces the risk of damaging something with the truck and increases efficiency. When loading with a hose, a good new position for the truck allows the hose to reach as many unfilled holes as possible. To determine the best position for the truck in a new row of holes, the hose is detached from the truck and used to measure out the ideal location. The output end of the hose is placed in the hole furthest away in the new row of holes. The part of the hose that connects to the truck is then placed as far back in the direction where the truck is coming from. The end of the hose that attaches to the truck is a beacon for the truck driver and spotter for where to stop the truck.

Using knots to indicate hole status

If a hole is wet, there is water at the bottom of the hole. Some types of product can push any water out of the hole and this works best if the product is released below the water. Because of this, it is important that the hose is dropped until it reashes the bottom of a hole. As the product gets released under the water, the water will be pushed out at the top of the hole.

When baiting the holes, crewmembers use different knots for wet and dry holes when they tie the detonation cord to the pole next to the hole. These knots signal to the hose operator that a hole is wet.

Stone as a fault warning

After baiting a hole, the booster should reach the bottom of the hole. While lowering a booster into a hole, the booster can get blocked and will not go down by itself. If the crew is really unlucky, the booster gets stuck in such a way that it cannot be pulled up either with a soft pull. This is when crewmembers need to be extra careful. Pulling too hard on the cord can cause the cord to snap, set off the booster and create an explosion. If a booster is stuck like this, the crewmember baiting puts a stone on top of the hole. The stone is a reminder to the crewmember at which hole this happened, a signal to other crewmembers something is wrong with the hole, as well as a blockade to anyone who wants to work with the hole.

3.7 Considerations of everyday work observations

Our observations generated a number of questions:

- "Are these actions good for safety?"
- "Are they desirable or necessary?"
- "Are there alternatives to these actions?"
- "Where did these actions come from, if not from the training or procedures? "

These are only some of the questions that can be asked, but they do show questions can go in different directions. Some questions are about understanding what is happening, some are about how to judge these actions, and others look for alternatives. Some questions look at the activity and team performing them, while other questions look for relations with other aspects. Depending on the view on safety and goals of the researchers some questions get more or less priority, and some questions might make no sense at all.

There are many differences between different schools of thought and some of the differences are very subtle. We have, however, aimed to capture the considerations that set different schools of thought on safety apart the most. By making these considerations explicit, we aimed to help mould everyday work investigations into something practical.

3.7.1 Goal

The first consideration – which will shape all other considerations – is the goal of the investigation. An investigation is either; 1) normative, when it tries to find the best way of doing things or find where things are currently suboptimal, 2) descriptive, when the goal is to describe and understand what currently is happening, or 3) formative, when there is a search for new possibilities to do things.

Normative approach

An investigation with a normative approach has the goal to prescribe how work should or should not be done, or how a system should or should not behave. In normative investigations observations are compared and can be benchmarked. This is either to find a norm, or work is compared to a predetermined norm. A normative question we could ask for our blast crews is "Are these practices good for safety?"

Norm-finding

In norm-finding normative investigations, the goal is to find the best way to do a job on one or multiple measures, which then are turned into instructions or constraints for the workers. The first task analyses from Taylor and Gilbreth did just this (Annett, 2000). Instruction-based approaches provide guidance for the workers what to do, should lead to workers do a task in the best-known way, lead to less variability between workers and should prevent workers from making errors. In closed systems, where the context tends to be similar, this works well. In open systems it is difficult to say in advance what a worker should do, as the context changes more often, which makes it difficult to give a single set of instructions. An alternative for these systems the constraints-based approach. In a constraints-based approach the goal is not to tell the worker how the job should be done; rather the goal is to tell what a worker cannot do. By only providing constraints, workers have freedom to manoeuvre and adapt to the circumstances as they see fit within those boundaries (Vicente, 1999). For our blast crew study we could look for different practices from other crews and compare these practices to the examples we identified. If certain practices are linked to more desirable events, than these practices could be used as instructions for other crews. If a certain practice is linked to undesirable events, constraints could be considered against such practice.

An example of a norm-finding study in troubleshooting comes from Schaafstal, Schraagen, and Berlo (2000). The researchers were asked to assist the Royal Dutch Navy in improving the individual skills of weapon engineers in diagnosing and fixing equipment problems. The researchers began by searching for expert troubleshooting strategies – both in the literature, and by testing and observing technicians of varying ability. As is typical for a norm-finding study, the goal was to find one or more “best ways” to undertake troubleshooting, which could be shared from the experts to the entire cohort of weapon engineers. Based on their investigation, the researchers designed a new training course around functional understanding of systems and a systematic search for faults. Originally this training was tested as an add-on to the existing training for the technicians. However, when skills of the trainees improved in subsequent tests, an entire training program for the engineers was redesigned around the new ideas. This redesigned training course was shorter than the original training, but when the trainees from this course were tested, they solved more problems in less time per problem than those trained the original way.

Other examples of norm-finding studies include Sanda et al. (2014) about drilling activities in mining, Rasker, Post, and Schraagen (2000) with feedback styles in command and control centres, and Staszewski and Davison (2000) on mine detection strategies.

Norm-comparison

Norm comparison is about how a system performs compared to pre-set norms. This is a common element in safety management already, it includes audits and compliance monitoring, which are considered the standard practice (Hopkins, 2007). However, when audits are used as a tool to change behaviour directly, rather than to learn about behaviour, they would not fall under our definition of everyday work investigations. With this approach, it is easy to make sense of what is observed and to determine what needs to be achieved. A norm comparison study highlights the current state in relation to a norm, but it is limited in that it tells nothing about the norm that is used. Even if it is well established that a certain way of doing a job is linked to fewer slips, it might have other undesired side effects, such as fatiguing people more or being less resource efficient. Norm comparison studies are not sensitive to unintended consequences or possible good reasons why people deviate from a norm.

A ‘soft’ normative approach avoids judging whether the difference is good or bad. This approach builds upon theories made from accident investigations that very much tried to understand what everyday work was like before the accident. Both practical drift (Snook, 2000) and normalisation of deviance (Vaughan, 1999) see deviations from standard procedures as both inevitable and necessary. Investigation of work, not instigated by accidents, confirm that deviations and work arounds are often necessary (Kontogiannis & Malakis, 2012). Over time however, deviations in multiple areas can interact with each other, as well as practices drifting beyond the boundary of safe operations. Therefore it is important for organisations to investigate their practices and reflect on whether they are safe (Dekker & Pruchnicki, 2013).

Our initial observations of blast crews walked the line between descriptive and norm comparison. We aimed to capture what was going on, i.e. a more descriptive approach, however, as we were surprised by what we found, and thus we started comparing it to norms in the form of procedures. This added norm-comparison elements. However, after finding no relevant link to procedures, continuing with a norm comparison approach would have been impossible, because there was no “norm” against which to compare. A common critique on many normative investigations is not necessarily that they are normative per se, but that they make unrealistic assumptions about human work. The way these models assume work should happen is often so different from how work actually happens that these models lose their usefulness (Sheridan, 2001). This critique applies more to instruction-based and norm-comparison approaches than constraint-based and soft-normative approaches.

An example of a norm-comparison study comes from Rodrigues et al. (2015). They looked into whether safety climate affected safety performance and risk acceptance across multiple furniture companies. The study was norm-comparative in that it started with a pre-established checklist for safe conditions, behaviours and procedures. The goal was not to investigate or verify these norms but to find out what caused organisations to move towards or away from the norms—specifically, whether better safety climate correlated with closeness to the norms. A questionnaire was used to measure safety climate and risk acceptance. To measure safety performance an audit was used, which is a norm-comparison approach. They had a pre-established checklist with 112 items about on the conditions of the workplace, equipment and machinery, and safety behaviours and procedures. For the audit, the researchers rated organisations on each item from 1 to 5 for how well they were doing. The pre-established items show that the researchers already had an ideal picture of work. As a normative study, the study needs a way to compare and make sense out of the data. This study, like many quantitative studies, used a statistical test to look for correlations and to see whether random variation could explain the found differences. The test found that high safety climate was linked with high safety performance and low risk acceptance. From this, the researchers concluded that safety climate influences safety performance and risk acceptance, establishing high safety climate as an additional norm to strive for.

Other studies doing norm-comparision investigations include Rydenfält et al. (2013) with compliance to the World Health Organisation surgical safety checklists and Garnerin et al. (2008) compliance with protocol on patient identity.

Descriptive approach

In a descriptive study a researcher looks at how a job is done or how a system behaves and why, opposed to how it should behave. Descriptive investigations do not immediately provide clues on what to do, but can help reframe problems, help uncover interesting areas, or provide a general understanding, from which decisions can be made. In our example of blast crews, the investigation in which we identified the practices was descriptive. Without this we could not explain what was happening. Follow up questions such as 'where do these practices come from?' fit with a descriptive approach. A field that advocated a more descriptive approach is Naturalistic Decision Making (NDM) (G. A. Klein, 1998; Orasanu & Connolly, 1993). NDM is interested in describing the process with which people made decisions (Nemeth & Klein, 2010). In the US descriptive approaches to work became popular with NDM and distributed cognition, while in Europe descriptive studies on work have been around longer (de Keyser & Piette, 1970; Rasmussen & Jensen, 1974).

A limitation of descriptive studies is that they only look at the current ways of doing things, and do not explore other possibilities. Descriptive studies do not examine how an intervention would effect the described processes (Woods & Dekker, 2000). It can also lead a descriptive study to focus on a 'workaround', while with a small change this workaround would not be needed at all (Vicente, 1999). The lack of focus on a solution means the results can be applicable to wide set of decisions, but can make it difficult to translate it into concrete interventions.

Carvalho et al. (2006) provide an example of a descriptive study, with their investigation of cultural and cognitive issues of nuclear power plant operations. They applied an ethnographic approach, using observations during work and simulations, and post-event interviews with control room operators. The study provides in-depth descriptions of how crews interact with each other and with technology, and how they deal with certain events. From there, the researchers focus on particular issues such how distractions affect operators and in turn how the operators deal with distractions. Whilst the researchers do ultimately make suggestions for improvement based on their results, improvement is not the focus of the study and did not guide the research design.

Other examples of descriptive studies include Carim et al. (2016) on how pilots use procedures in non-normal situations, Schraagen (2011) on coordination and communication in medical team during non-routine situations, and Vicente, Mumaw, and Roth (2004) on nuclear power operators during normal operations.

Formative approach

When taking a formative approach the goal is to find new ways of doing things. Vicente (1999) defined a formative approach as "Based on identifying requirements that must be satisfied so that a system can behave in a new, but still desired, way". Knowing more ways of doing things increases the flexibility of the system (Vicente, 1999) and can be a useful goal in itself. While many research projects have the end goal to find a new and better way of doing things, there are few methods that are formative. Cognitive Work Analysis (Vicente, 1999) and ethnography (Dekker, Nyce, & Hoffman, 2003) have been suggested to be able to fulfil this role, by looking at what a system does and why it does what it does. By understanding the goals and constraints that influence a system, alternatives can be sought that fit with these goals and constraints. However these methods do not have a structured process or step-by-step guides. Theoretically Cognitive Work Analysis has clean structured process with analyses on multiple levels. However, in practice the methodology is often messy and analyses on one level find results for another level (Vicente, 1999).

A detailed example of an entire formative analysis comes from Jatobá et al. (2016). In this study, the researchers aimed to understand and improve patient risk assessment process used for prioritising patients. Their main question was "How can work situations be enhanced and support devices be designed in order to improve the risk assessment process in primary health care?". The study does build a somewhat descriptive understanding of what is currently done, but the description is in support of finding new ways it is not a goal in itself. The formative intent influenced the way the study was conducted from the very first step. Jatobá et al. began by 'framing' the entire study, to understand what the managers and workers wanted to improve and sought to achieve with the improvement process. Next, they used observations and interviews to evaluate various features of the workplace. From there, they then observed and modelled the work processes, identifying specific problems and opportunities for improvement. The study did not compare the work pre-established model like a normative investigation would. Unlike a purely descriptive study, Jatobá et al. constantly engaged on what was imperfect and could be improved.

Not all formative investigations need to be as elaborate and formative from the beginning. With our blast crews, a question such as "Are there ways other than the observed ways of doing things?" falls into the category of formative studies. Cornelissen et al. (2012) and Hassall and Sanderson (2014) describe methods to search for alternative ways work can be done. Both start with an investigation into the existing domain, goals, constraints and strategies that make up work. The step where investigation into strategies becomes formative is where they differ. Cornelissen et al. (2012) tries to link existing actions to different parts of the system, to find new combinations of goals and strategies. If we apply this to our blast crews, we could look into whether objects other than hoses could be used as beacon. For our blast crews using hoses as a beacon, using this method, we can think of alternative strategies to determine where to park the loading truck, like 'intuitive' strategies where crew members use a 'fingerspitzengefuhl', or crews that use an 'analytical' strategy, and calculate a good position for the truck based on the number of holes, the distance between the holes and hose length. Both the method of Cornelissen et al. (2012) and Hassall and Sanderson (2014) are formative in that they are systematic and go beyond describing how things are, and look for other possibilities.

Formative investigations are successful when they find new ways of doing things that work, but it is difficult to know how valid or complete an investigation is. The results cannot be compared to anything. A quick-and-dirty formative study might find many options that only later shown to be poor alternatives, while a thorough study might fail suggest any changes that it considers desirable. No matter which formative method is used, no kind of analysis can find all possible ways of doing things(Cornelissen, McClure, Salmon, & Stanton, 2014).

Discussion

The lines between the goals above can become blurry. Studies that set out to be normative can expose new ways in which people are doing things. Descriptive studies are often part of a larger project with normative or formative goals. Even if it is not part of a larger project, descriptive studies often suggest how a system can be changed for the better. Formative studies often rely on descriptive and normative methods.

3.7.2 Post-positivist versus social constructivist view

For any investigation of everyday work there is the question about how we view knowledge about safety. In a post positivist worldview we want to know about something that exists in an objective world outside of people's subjective perception (as applied by, for example, Normal Accident Theory, the Swiss Cheese Model and Behaviour Based Safety). If safety is approached in a post positivist way, it is either something a system has or something it does. Definitions of safety as "freedom from unacceptable risks" (Lowrance, 1976) suggest that safety reflects something that exists in an objective world. The line between acceptable and unacceptable might be subjective, but the risks and freedom from it exist in the world. Investigations with a normative goal tend to have a post positivist view of safety. They believe there is something that can be done better. One can still recognise that safety means different things to different people, but this is then a matter of semantics. There is an underlying objective goal to be reached by improving safety, which is free from subjective perception. In safety this is usually related to decreasing the number of accidents, damages, fatalities or injuries.

Another social constructivist view, in contrast, considers what and how meaning is given to something by people. Studies that take this approach tend to be descriptive. The early High Reliability Organisations studies are an example (Rochlin, 1999). The researchers found that safe operations did not just mean not having accidents for practitioners; it was about what was done to prevent accidents. This was a lot more than avoiding errors; this also included being critical of current operations, learning from near misses and looking ahead to what might happen in the future. This changes which activities relate to safety and how safety can be improved. To relate this to our blast crew observations, it raises the question what value crewmembers see in the actions we observed. What part do these activities play in their view and what value is assigned to them? The crews might not only see the knots as communicating information about the hole, but also as a way to stay mindful, to use a High Reliability Theory concept. Mindfulness in turn can be an important piece in how reliability or safety is created (Weick et al., 1999).

There are more worldviews than positivism and constructivism, but most fall on a spectrum between these two. Resilience Engineering, for example, has both been used from a post-positivist perspective and social constructivist perspective (Le Coze & Pettersen, 2008). Positivism has the advantage that it is easier to combine with hard sciences, as they are often rooted in a positivist worldview, and the studies often provide clear goals to aim for. It can be harder to combine with the understanding of multiple diverse individuals and the concepts the study uses can function as a lens that make it blind to things that not fit within these concepts. Constructivism is more sensitive to diverse human viewpoints and the study of things that are difficult to define in advance, but as it lacks a universal vision of truth or good, and thus can be difficult to translate in interventions.

3.7.3 Level and units of analysis

The level of analysis refers to the scope of what is investigated: how micro or macro the orientation of a study is. What is significant at one level of analysis is not necessarily so at another level of analysis. A study can look at multiple levels of analysis, this, however, comes with extra challenges (K. J. Klein & Kozlowski, 2000). Most schools of thought in safety do not deny the relevance of different levels of analysis in general, but can place more focus on some levels, or give their advice on a specific level. Normal Accident Theory and High reliability Organisation work both place the emphasis on the organisational level, but do not rule out that the actions of individuals matter. In the High Reliability Organisation literature, the actions of one individual are described, but the meaning that is given to these actions is about how this fits into the organisational processes (Weick et al., 1999). The latent failure model says actions on the individual level always play a role, but looks how other levels influence these actions. Resilience Engineering can be applied at different levels and across multiple levels, as it sees differences in goals across levels as a reason systems can fail (Woods & Branlat, 2011).

The unit of analysis is about the type of things an investigation looks at. Safety research often used accidents as a starting point, but everyday work analysis moves away from this. One way adjust is to rely on what accident models describe as relevant. Normal Accident Theory would tell investigations to focus on qualities or features of an organisation. Heinreich's accident pyramid would guide you towards incidents and unsafe acts, which returns in behaviour based safety (Krause, 1997). The latent failure model would talk about barriers and holes in them. However, applying these concepts to everyday working, without the hindsight of an accident, can be challenging. The schools of thought that focus more on everyday work provide easier to apply concepts. High Reliability Theory looks at the processes with which people organise in risky fields (La Porte & Consolini, 1991). Resilience Engineering often looks at what it takes to succeed in varying conditions and with that often ends up looking at adaptation (Hollnagel et al., 2006). Even in fields that study people at work, there are still (subtle) differences. For example, Naturalistic Decision Making and Cognitive Systems Engineering both look at people at work and could design a similar study. However Naturalistic Decision Making tries to analyse how people make decisions, and the models it produces are about what happens (mostly) inside their head, while cognitive systems engineering leans more towards analysing what happens at the intersection between people, technology, and work, leaving it more of a black box of what happens inside the head of people (Hoffman & Militello, 2008; Hollnagel & Woods, 2005). Studies of everyday work have often proven to be useful for the improvement of safety, but it is not a given that they can be used for safety. Overall, this shows the challenge of finding relevant unit of analysis for investigations of everyday work.

In our investigation with blast crews we started looking at the team level and interactions between team members, simply because it was the most obvious way in which work got done and aligned with the exploration of possibilities for Crew Resource Management training. If we ask the question “where did these practices come from?” we are likely to look at other levels of analysis, such as the influences of individuals or organisational features. If we have hypotheses on what form these practices, such as mine policies regarding stones on benches, we could do a normative study and test whether these practices are more likely to develop in the hypothesised conditions. If we have no hypotheses, we are more likely to take a descriptive approach and find how a practice comes to be. Levels of analysis do not only apply to units of analysis that are systems or objects. Research that focuses on actions, like organisational routines, treats actions as subsets of other actions. Interviewing is an activity that is nested within the activity of hiring a person (Pentland & Hærem, 2015).

In most investigations the unit of observation is close to the unit of analysis. If the goal is to know about whether people wear personal protective equipment (PPE), the investigation will probably look at people wearing their PPE in a number of situations. However, many units of analysis are infeasible or impossible to observe directly. The tidiness of a site, for example, is observable, but it can be assessed in different ways. Even when people agree on its definition, there are different ways to operationalise it. One can investigate what tools are not in their predefined place at the end of the day, how long it takes before tools are found, or count objects in an area that are not supposed to be there. The unit of observation is about how the unit of analysis is operationalised, which includes both objects being looked at and when it is being looked at.

In practice, there are usually multiple units of analysis and each unit of analysis can have multiple units of observations. For work design, Carayon et al. (2007) suggest the units of analysis the organisation, the tools, the environment, the tasks and the individuals, as a starting point for work design to improve patient safety. The units of analysis go across multiple levels of analysis and for each unit of analysis they list multiple analysis techniques that come with different units of observation.

3.8 Conclusion

In this paper, we have used our empirical observations of blast crews to both inform practical steps for doing everyday work investigations with a focus on safety, and reflect on their meaning vis-à-vis various schools of thought.

In the different choices there is a tension between 'for safety', the prevention of specific special events, and the 'everyday', something unspecific and ubiquitous. A normative and positivist approach, which uses units of analysis based on existing theory, is theoretically clearly relevant for safety. However, it is prone to preconceptions that poorly fit with the everyday work that is being investigated, which hampers learning potential. A descriptive and constructivist approach offers more possibility to investigate everyday work and to learn things that are unlikely to be discovered through a normative, accident-prevention focussed approach. However, the link between the research and 'for safety' is more indirect and it will often be harder to translate the findings into interventions.

For this reason, there is a trend for safety-motivated everyday work research to reduce emphasis on safety as an end-goal. Safety-II focuses on what helps create success, and recognises that not everything investigated will turn out be relevant for safety. (Hollnagel, 2014). Both Resilience Engineering and High Reliability Organisations do not talk about safety, but rather "resilience" and "reliability". These terms are safety-adjacent, and but they are not the same as safety. At best they consider safety as one consideration in determining whether a successful outcome is achieved (Leveson, 2011; Morel et al., 2008; Weick, 1987). The shift from 'safety' to 'success' means that everyday work investigations may be inappropriate for considering hazards that do not get in the way of short-term success. For example a worker may experience long-term harm (such as from asbestos) but may appear to be working successfully.

By reflecting on research practice, we have uncovered and discussed assumptions that are shaping everyday work investigations. Critical consideration of these assumptions will allow practitioners and researchers to conduct everyday work investigations that open up access to new knowledge whilst retaining relevance for safety.

3.9 Outcome for the research aims

Based on these reflections, I decided that my research of blast crews should be descriptive. Similar to the early HRO work, my goal was to understand how crews organise for safety. In the CRM review, I discovered that there was no norm to compare blast crews to, ruling out a norm comparison study. There were no indications that some crews were better at organising than other crews, nor did I have measures to compare "quality" of organising on, ruling out norm finding investigations. My research aim of how teams organise was asked with a desire to improve, but the goal is first to understand. Improvement opportunities also do not need to come from the organisation of the teams themselves, but could be achieved by adjusting what feeds in crews or what constraints teams. This is why descriptive research approach was chosen. An understanding of how things are organised is useful as it can feed into decisions in an organisation in multiple ways and inform future interventions.

I looked at what groups did and how they organised. I had identified that blast crews used signs to communicate, however, I did not know the details on what these signs meant or how they came about. To understand how blast crews organise, I needed to understand what these elements meant to people in blast crews and what their relevance was. This need directed me to a constructivist approach.

Our level of analysis was first and foremost the team level, as per my research aims. Individual elements and organisational elements were taken into account, but this was done to explain how things played out at the team level.

My unit of analysis was the organisation of crews, but I decided to limit the scope to the organisation of routine work. Non-routine work was still worthwhile to investigate, as it could show what was different about it compared to routine work, but not as a subject to analyse by itself. The interest in routine work came from the observation that blast crews coordinate their work with surprisingly little verbal communication. When I asked blast crew operators about how people know what to do, without a stable role distribution or explicit coordination, the short answer usually was that they were all experts, had done the job a million times, and knew what to do. This explanation can be understood that it is general expertise with the task, but this explanation does not fit perfectly, as there were predictable differences between teams in how they did things. The organisational routines literature could explain the non-verbal coordination with the repetition in actions, but this explanation could also be doubted, as blast crews still adapted to each other and the situation. It could not all be just repetition. With an investigation into routine work, I aimed to incorporate both the repetition and the adaptation.

Chapter 4: **Organising routine work**

4.1 Statement of contribution to co-authored submitted paper

This chapter includes a co-authored paper. The bibliographic status of the co-authored paper, including all authors, is:

Havinga, J., Dekker, S., & Rae, A. (2017). Routine variation: How routines facilitate adaptation. *Manuscript submitted to Cognition Technology & Work*

My contribution to the paper involved:

The design of the field work, with normal PhD supervision input from Prof Dekker and Dr Rae; The ethics of the study was jointly considered with Prof Dekker, & Dr Rae; The ethics application; The planning of the field work; The field work and analysis, with informal discussions with Dr Rae on the topic; The writing of the first draft of the paper; Minor edits and elaboration based on feedbacks from Prof Dekker and Dr Rae.

(Signed) ______________________________ (Date)____________

Joseph Paulus Havinga (Corresponding author)

(Countersigned) ________________________ (Date)____________

Dr Andrew Rae (Principal supervisor and co-author)

(Countersigned) ________________________ (Date)____________

Prof Sidney Dekker (Principal supervisor and co-author)

This study was conducted under ethics registration GU Ref No: 2015/191

4.2 Rationale for the study

While the idea of investigating teams at work fits with HRO and Resilience Engineering, the idea that teams rely on routines and repetition is at unease with both, or is at least under-addressed in both. Neither HRO nor Resilience Engineering, argue it is impossible that teams rely on routines or repetition to organise their work successfully, but both emphasise reliance on its opposite, adaptation and variation.

HRO mentions that organisations keep their outputs reliable and stable by varying their processes that produces that output as they adjust to a situation. The idea of a routine, or standard way of doing the job is sometimes even portrayed as the enemy of mindfulness (Weick & Roberts, 1993; Weick et al., 1999), which is the awareness of operations needed to deal with the unexpected. In that way, according to HRO routine action does not fit with the goal of creating reliability.

Almost the same argument applies for Resilience Engineering, although there is more acknowledgement of the general existence of routines in (reliable/safe) organisations. Resilience Engineering tends to stress the adaptations required to fill the gaps in the current system (Woods & Wreathall, 2008) or the imagined system. Resilience Engineering and CSE acknowledge that all systems have some kind of path-dependency or returning patterns with which they address problems. For example; systems must constantly adapt to their environment, but part of the system must always change slower than the environment for the system to sustain itself (Cilliers, 2005), that all complex systems have a path-dependency, that influences current behaviour (Dekker, Cilliers, & Hofmeyr, 2011), and that all systems are already pre-excited and organised and there is no such thing as a static system (Hollnagel & Woods, 2005). However, I could find little explanation of what a routine is or what it means for a team to follow a routine. Nor could I find that repetitions are required, or relied upon, in how work is organised.

This lack of focus on what repeats turned me to the organisational routines literature. As mentioned, the organisational routine literature could not explain my observations in detail either, as it could not explain the variations I observed. However, organisational routines theories provided a starting point and something to compare my observations to.

By just looking at blast crews, it would be hard to make claims about the nature of routines in general. To enhance the external validity of the study, I have also investigated power line crews of an electricity distribution company. Power line crews made a good comparison group, as they also have some stability in team composition and bring their own tools to different environments.

Routine variation: How routines facilitate adaptation

4.3 Abstract

In this paper we report on the results from a problem-orientated ethnography with blast crews and power line crews. The investigation studied the nature of routines, and how routines assist in the coordination of work within a team. Extending theory from Cognitive Systems Engineering and organisational routines, we provide an explanation for how individual decisions give rise to general repeated patterns of work at a team level. This explanation suggests that when teams work together, their experience leads to a shared problem understanding and solution structures, which give rise to the signs that coordinate work. These structures provide stability in how crews complete a task and at the same time give freedom to adapt within the structure, and can even enable adaptive use of signs. It is through these structures that past work influences present work. Our explanation questions the distinction between heedful interrelating and routine interactions, and accounts for the observation that experts in teams become more flexible, rather than more fixed as some theories on routines imply.

4.4 Introduction

In dynamic systems, people and teams continuously adapt to evolving situations, but there is also path-dependency, i.e. their behaviour is influenced by what they have done in the past. Part of this path-dependency takes the form of standard practices or team routines.

Existing theories of individual routines are based on the notion that people are guided by ideal patterns of actions. The patterns of actions develop with experience so that greater expertise involves more fixed behaviour. Variation is described as a deliberate or accidental departure from routine that must be further explained.

Based on a problem-orientated ethnography with blast crews and power line crews, we propose a new explanation for what makes up a routine at a team level and how routines assist in the coordination of work within a team.

As teams acquire experience together, they develop shared understanding of the problem to be solved and the structure of available solutions. From their problem understanding and solution structure follow guiding goals and meaningful signs that co-ordinate their work process. It is the shared solution structure, rather than patterns of action, that repeats in a "routine". This explanation, supported by theories from Cognitive Systems Engineering and organisational routines, can account for how individuals can vary their own behaviour but also give rise to general repeating patterns at a team level.

Our explanation contradicts the notion that people are guided by ideal patterns of actions, or even that routines have a base pattern of action. It provides a mechanism for understanding how teams coalesce, and is consistent with the observation that experts within teams become more flexible, rather than more fixed, as some theories on individual routines predict.

4.5 Literature review

4.5.1 Individual decision making

There are multiple theories that describe the mechanisms of routine activity and decision-making for individuals. Operant conditioning describes how organisms associate stimuli with actions, and suggests that this association forms the most basic building-block of learning (Skinner, 1938). The skills-rules-knowledge framework (Rasmussen, 1983) contests this idea, proposing a hierarchy of decision-making mechanisms. Slow knowledge-based reasoning can be replaced by faster rule-based and skill-based decisions as expertise develops. In this process, computational demands are traded for perceptual demands (J. M. Flach & Voorhorst, 2016). Skill-based and rule-based decisions are more complex than stimulus-response, because they are directed towards a goal. Other decision-making theories suggest that schemas are used to link current situations to associated experiences to avoid complex computations. Schemas are also used to interpret actions of fellow crew members (G. A. Klein, Orasanu, Calderwood, & Zsambok, 1993). On an individual level, these ideas provide a foundation for understand routines, but are insufficient to explain the phenomenon of team level routines.

4.5.2 Multi-actor routines

Routines have been used to account for the functioning and malfunctioning of socio-technical systems. Klein, Feltovich, Bradshaw, & Woods (2005) say that the routines a team can perform are part of the common ground between members, and so provide a foundation for the coordination of joint activity. Woods and Branlat (2011) also mention that getting stuck in outdated behaviours is a pattern in which adaptive systems fail. For example, according to Weick (1990), the behaviour of one of the captains in the Teneriffe accident leading up to the accident was a reflection of older learned patterns. Getting stuck is not the only failure pattern for routines. Snook (2000), Vaughan (1996), and Dekker (2011) talk about how groups change their routines over time, and argue that this unavoidable process can lead to reduced safety margins and occasionally to accidents. All of these ideas acknowledge the importance of path-dependency – the way in which present behaviour is shaped by past behaviour. However, beyond using the term “routine” as a general encapsulation of path-dependency, none of them explain exactly what a routine is.

4.5.3 Organisational routines

The concept of organisational routines, coming from organisational studies, explores in greater depth the path-dependent behaviours performed by multiple actors. The phrase “organisational routines” was introduced by Nelson and Winter (1985), but reflected established ideas that organisations often operate by matching known procedures or scripts to known problems, rather than by systematically planning and calculating the best course of action (Cyert & March, 1963; March & Simon, 1958).

As organisational routines were introduced to describe a lack of planning, they were often seen as something mindless. "Routine" was synonymous with both low-skill work and inertia. This view began to shift when many empirical studies found that variations of routines were the norm, not the exception, and that change was inherent to routines (Feldman, 2000; Pentland & Rueter, 1994). Organisational routines have subsequently been observed in many types of work that are not considered mindless or low skill work at all, including medical work, engineering, legal work, and research & development (Pentland & Hærem, 2015). Instead of being viewed as mindless repetition, organisational routines are now seen as both a cause for stability and a mechanism of change (Feldman & Pentland, 2003; Pentland & Hærem, 2015).

A key point in the turnaround from mindless to effortful was the conceptualisation paper of Feldman and Pentland (2003). They defined organisational routines as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (p.95). This definition resembled older definitions, however, Feldman and Pentland also introduced a new ontology. They proposed that organisational routines were made of a duality of an 'ostensive' and a 'performative' aspect.

The ostensive aspect of a routine is the ideal image of a routine—a summary of how things are usually done. The ostensive aspect is what agents recognise as the general repeated pattern. This is not necessarily a single unified image; different agents can create different images of the same activity. The ostensive aspect is abstract and does not include specifics, as it needs to be applicable in various conditions. The ostensive is not the same as written rules or procedures, as those are artefacts, however, artefacts can be based on the ostensive routine (Feldman & Pentland, 2008; Reynaud, 2005).

The performative aspect is the enactment of a routine. This includes who performs the actions, the time, the context, and the specifics of how the actions are done. The performative aspect is always improvisational and includes deviations from the ostensive abstract image.

The ostensive and performative together make up the routine, and continually influence each other. The ostensive aspect *guides* practitioners like a script when they are performing a routine. The ostensive aspect allows practitioners to *account* for what they are doing and helps them make sense of what is going on. The ostensive aspect also allows practitioners to *refer* to the performative in order to plan activity. The performative aspect *creates* the ostensive aspect as practice happens, and *maintains* the ostensive part through repetition. Through the variations in the performative aspects, the performative aspect *modifies* the ostensive aspect.

With the ostensive and performative constantly influencing each other, Feldman & Pentland (2003) could explain both the variation and consistency observed of people at work. With stable teams, the team can start to coalesce around a naturally followed action sequence, where each member of the team knows what to do and what to expect (Turner & Rindova, 2012). This is not to say that routines happen automatically. It requires skill and effort to perform a routine in accordance with the ostensive aspect under varying conditions.

In a recent review, Pentland and Hærem (2015) explore the consequences of this conceptualisation of routines. They say that routines have their own level of analysis, or granularity. For example, an interviewing routine can fit within a hiring routine. Unlike actor-based units of analysis, routines can be infinitely far decomposed. Pentland and Hærem also mention that this model of routines does not explain how actions in a routine link together.

4.5.4 Heedful interrelating

Weick and Roberts (1993) introduced the idea of heedful interrelating as the opposite of routine interaction. A heedful actor goes beyond local concerns to take into account larger task and organisational factors when interacting with others. Heedful interactions are modified by each replication, contrasted with routine interaction, which just replicates what has been done before. According to Weick and Roberts (1993), heedful interrelating prevents accidents. Based on their study of carrier flight-decks, they concluded that the rotation of personnel across teams keeps people in a mode of heedful interrelating, preventing routines from forming.

A few studies have tried to operationalise heedful interrelating and have found mixed results. Grote, Kolbe, & Künzle (2010) analysed the behaviour of multiple cockpit crews in a simulator. They did not find that high performing cockpit crews exhibited more heedful interrelating, but they did point to qualitative differences in how the crews heedfully interrelated. They also found that heedful interrelating was more common during high workload tasks, and less common during more standardised tasks. In addition, their results showed that implicit coordination was higher in both high workload and high standardisation conditions. These results support Weick and Roberts' (1993) suggestion that heedful interrelating is different from standardised or routine action, but they diverge from Weick and Roberts (1993) in that crews can still rely on 'routines' to coordinate non-standardised tasks. Grote et al. (2010) also link heedful interrelation to task characteristics rather than crew performance.

Schraagen (2011) analysed the interaction of one medical team during multiple paediatric cardiac surgeries and found that during non-routine events, the team mostly relied on explicit coordination, not heedful interrelation. However, in-depth qualitative analysis of the interactions suggested that heedful interrelations did make up critical moments. The study concluded that heedful interrelating is important, but that most non-routine events are dealt with through routine procedures. According to Schraagen (2011), routines provide building blocks to deal with events, even when the events are non-routine.

Gorman & Cooke (2010) found in experiments that the exchange of team members between crews does lead teams to search for more adaptive solutions, as was suggested by Weick and Roberts (1993). However, the study did not investigate whether this change was mediated by heedful interrelating.

To summarise, heedful interrelating has not been found to generally improve performance, nor has it been shown to be necessary for team performance during non-routine events.

4.6 Methods

4.6.1 Research setting

Blast crews

Blast crews' in mines use explosives to crush, or 'soften up', the ground. The process we will describe here reflects that which is found in coalmines, where the goal is to access the coal underneath the softened up ground. A similar process is used in quarries. If the blast crew's blast is too weak, the rocks afterwards are too big, making the job harder for the diggers in the mine. If their blast is too strong, environmental norms can be exceeded, and debris can be flung away from the blast area. Besides the risks posed by debris, the job of blast crews is safety critical due to the risk of premature and uncontrolled detonation, and the risks that come with operating and working alongside heavy machinery.

After an area to be blasted is surveyed, a 'shot' is designed. A shot is a plan that shows what kind of explosive 'product' to use and how the product will be distributed into the ground. The holes where the product will be loaded into are drilled per the plan. The blast crew checks whether the depth of the drilled holes are correct and, and puts pegs next to all the holes. Then, as they get ready to load, they will put the required detonation cords and boosters next to each hole. Boosters are small cylindrical explosives that are used to initiate the main explosive product. When loading product into a hole, the crew first 'primes' the booster, which means attaching the detonation cord to the boosters, after which they 'bait' the hole. Baiting means lowering one or multiple boosters attached to a detonation cord into the hole. At the same time, the product is loaded into the hole and the detonation wires are tied to the peg next to the hole. The hole is topped off with stemming. The last step is to connect the detonation cords together according to plan before the shot is set off from a safe distance.

Within blast crews, there are different roles and qualifications, such as heavy vehicle operating licenses and shot firer licenses. However a lot of work can be done by multiple people and is not delegated in advance. Instead, most of the work distribution arises on the spot, often without explicit verbal communication between the crewmembers. While crews can work for hours on a bench while hardly talking about work, their work is interdependent and builds on each other. Blast crew operators say they do not need to talk because of their expertise.

Our initial observation of crews at work showed that crews act upon a lot of physical signs, many of which are left purposely to communicate (Havinga, Dekker, & Rae, 2017). Signs include things like spray paint, traces of product, placed stones, and parked cars. What signs mean differs between crews, suggesting that are reflections of the path-dependency, or routine, of a crew.

The features and working conditions of blast crews provide a good opportunity to study routines. Blast crews usually have the same team members, which gives them the time and stability to invent their own practices. Much of the work, as well as the communication, has a physical element to it, which makes it easier to observe and study. On top of that, mines are often located far from each other and form relatively closed systems, with limited interactions between crews from different sites. This makes it easier to trace outside influences. Between mines the tasks are often similar, but the conditions differ, both organisationally and environmentally. These features and conditions make blast crews an extreme case for the study of routines (Flyvbjerg, 2006), as it displays the ideal conditions in which different routines are expected to develop.

Power line crews

In addition to blast crews, we also studied power line crews. Power line crews have multiple jobs related to building, maintaining, and fixing the above ground part of the power grid. Their work includes placing new poles, connecting houses to the power net, and replacing crossarms on existing poles. As the main jobs of power line crews are diverse and do not form a sequence, we do not go into as much detail.

The work of lines men takes place in a less closed system than blast crews. Power line crews are organised by depots, each with their own districts. Within the depots there are some recognisable teams, but alternative setups happen frequently as does collaboration between depots. In addition, the work of line crews is in public areas, rather than in fenced off mines.

Like blast crews, power line crews bring their own equipment to each job. There are different qualifications lines men can hold, but a lot of tasks can be done by all crewmembers; meaning that not all tasks are delegated in advance. There is more talking and planning happening during the job, but still many activities are initiated without being explicitly planned or communicated. Compared to blast crews it is easier to predict who will do what, as crewmembers can less easily change positions when they are in a bucket truck, on a ladder, or on the ground.

4.6.2 Data collection

We studied both types of crews through problem-orientated ethnography. This included a field immersion for a total of seven weeks across different blast crews and power line crews. Workdays varied between 8-12 hour shifts, not including travel time spent with crews, during which data was gathered as well. The fieldwork consisted of observations and both unstructured and semi-structured interviews. Information was captured in notes, pictures, and audio recordings. Outside of the field immersion, training was followed, documents were studied, and some observations were verified through interviews with office-based staff within the organisations.

Researcher

For the study of blast crews and power line crews, the observer (the first author) underwent general safety inductions for the professions, as well as site-specific inductions. The observer was an outsider with limited understanding of the field. Because of the focus on actions, signs, and their repetition, we believe this novice status of the researcher was beneficial. Lacking the expertise of the practitioners that were studied, the researcher could not ascribe meaning to the environment in the same way the experts could. This forced the researcher to start questioning from observable details, rather than abstractions, which helped in making the mechanisms of coordination as explicit as possible.

Semi-structured interviews

The semi-structured interviews were based on cognitive task analysis methods (Crandall, Klein, & Hoffman, 2006). During these interviews we showed pictures of crews at work, and asked the interviewee(s) to answer questions in relation to the pictures. Each interview had between 1-3 practitioners. The main questions of the interview were:

- What is going on here?
- What are they (the crew in the pictures) going to do next?
- What would you be doing to help out?
- Is there anything that concerns you?

The interviews were semi-structured in that the interviewees were often prompted to go more in depth, to explain how they knew things and why they said something. In addition, the practitioners were asked about alternative ways to do things and asked about what would be done differently if certain elements were different. Practitioners were encouraged to ask for more information or ask for clarifications if things were not clear in the pictures. This in itself helped uncover what information was relevant for practitioners. If the interviewees' answers were far off from what was observed when the picture was taken, more information would be given about the work of that day, to see if this changed their answers. This enabled exploration of whether differences in answers reflected individual differences or differences in known information.

Practitioners were shown pictures of themselves at work, as well as pictures of other crews at work. When practitioners were shown pictures of themselves at work, the picture was mostly a tool to point to a situation, and the interview showed resemblance to a Critical Decision Method, in that it tried to uncover what cues were involved in making a decision (Crandall et al., 2006). Where people looked at pictures of other groups at work, it functioned more like elicitation by critiquing, as it would often show other possible methods and an exploration about the constraints of the work (Miller, Patterson, & Woods, 2005). The combination of the two functioned like an experiment-like task (Crandall et al., 2006), where the experimental manipulation was how much people knew about a situation, as well as the different routines practitioners had.

4.6.3 Analysis

We conducted a thematic analysis (Braun & Clarke, 2006), where elements of the data were coded and larger patterns were sought in the data. These were then tested through triangulation. The analysis commenced during and between data collection periods. This guided observations to verify earlier observations, to test initial hypotheses, and to explore recognised gaps.

Part of the analysis was data driven, but patterns that were actively sought were:

- The actions and order of actions with which work was conducted
- The information that operators based their actions on, including planning, communication, and meaning given to elements of the environment
- Differences and similarities of how work was conducted between multiple instances of a task:
    - Between multiple instances of one person performing a task
    - Between different operators in one crew
    - Between crews on one mine site or depot
    - Between crews from different mine sites or depots

- What crews did when they could not use their ‘go-to’ practices

4.7 Results

The result section is presented in three parts. The first part describes what varied (i.e. where no stable pattern was found). The second part describes routines, or what repeated between multiple instances of work. The third part describes how elements of a routine changed over time. In the discussion we combine these elements into one unified explanation.

4.7.1 Variation in actions

Neither blast crews nor power line crews showed repeating patterns of action that could be identified as a baseline, as has been suggested by the ostensive part of an organisational routine (Feldman & Pentland, 2008). Some observations could easily be classified as non-routine, such as when a phone call interrupted a practitioner, or other outside events interfered with the work process. There were events that might be explainable as routine variation from a routine, such as one person helping a struggling crew member out. However, this line of inquiry, at best made deviations visible; it did not show what pattern of actions was the norm or baseline. For many tasks, there were multiple actions patterns, within one crew, and even for one person. Closer investigation of the data pointed not to set patterns, but a more opportunistic interplay between the practitioner and the environment.

The case of measuring hole depth, as an example of the individual level, and the case of loading a hole, illustrate this interplay in blast crews. The explanation of a base ideal pattern from which is varied is not necessary for this explanation, nor does it fit for the diversity in actions observed.

Measuring hole depth

After the holes are drilled into a bench, a blast crew will measure the depth of the holes. This is to verify whether the holes have been drilled at the planned depth. During this process, teams will make note of whether there is water in the holes. For this task, the crewmembers use a dip tape. A dip tape is a long cord with length measures on it and a weight at the bottom. To measure the hole depth and water level, crewmembers lower the weight into the hole and feel the change in resistance as the weight hits the water surface and the bottom of the hole. Both the measured hole depth and the water level (measured from the bottom of the hole), will be written on a peg next to the hole. Table 9 describes the actions involved in measuring a hole depth, when there is water in the hole, and the peg is not in the ground yet.

Table 9 Hole depth measuring actions

Actions required for measuring hole depth
Walk to the hole
Pick up the peg
Stick the peg in the ground
Bring the dip tape weight to bottom of the hole & read the dip tape
Write the hole depth on the peg
Bring dip tape to the water level & read the dip tape
Write the water level on the peg

These actions did have a set pattern, or a baseline. Picking up the peg and sticking it in the ground did not always happen before the dip tape was put in the hole. Sometimes the peg was done first, sometimes later, and sometimes the peg was picked up while the dip tape was lowered. Sometimes the hole depth and water level measuring were conducted in the order of Table 9, but sometimes both were measured in one go, and written on the peg afterwards.

These differences could not be explained as variations from a base pattern. The differences in order of the actions were not necessary; any of the observed patterns would have worked in any of the situations. Neither did the variation reflect individual preferences, as most operators in the crew showed all patterns. The thing that could explain the differences in actions was what the environment presented to the crewmembers first. If the peg was on the path towards the hole, the crewmember would first pick up the peg and then start measuring, while if the peg was behind the hole, the crewmember would start putting the dip tape in the hole, and then walk over to pick up the peg. There was no standard location of the peg in relation to the hole, and so no baseline pattern.

The order of measuring of the water level and hole depth and writing this on the peg, could be explained by the water level. If the water level was low, crewmembers would write both the hole depth and water level at once, while if the water level was high the crewmembers would first write down the hole depth and then measure the water level. This was, however, not a simple if-then rule, as when they lower the dip tape and hit the water surface, the practitioners cannot know the water level yet, as an operator needs the hole depth to calculate this. Instead, with low water levels, the operators have the experience of the water level still in their working memory. This means they know the information needed to write the hole depth and water level on the peg, their larger goal, is within reach. An outside view can in hindsight, recognise different patterns of actions for different conditions, but from a personal perspective, people are not following a script of actions based on if-then rules, but rather using what they see as available towards their goals.

Loading a hole

When blast crews load a hole from a truck through an auger, there is always one person operating the truck, with one, two, and sometimes three people next to the truck on the bench. These will be called bench hands. Crews can have tendencies in the number of bench hands they use next to the truck, but with any number of bench hands, the pattern of actions varies. Table 10 shows the (minimum) actions needed to load a hole that needs two boosters.

Table 10 Loading a hole actions

Action required to load a hole
Truck drives over to hole
Bench hand walks over to the hole
Bench hand primes 1st booster
Truck loads bottom bit of product into the hole
Bench hand 1st booster is lowered into the hole
Truck loads main load of product into the hole
Bench hand checks height of product
Bench hand 2nd booster primed
Bench hand 2nd booster lowered into hole
Truck loads last bit of product into the hole
Bench hand checks product height

This set of actions has less freedom to change the order of the actions, without messing up the end result. A booster cannot be lowered into the hole without it being primed and the main load cannot be loaded before the first booster is in. What can be varied is the moment when the boosters are primed in relation to the loading. The first booster can be primed before the bottom load, or after the load is already in the hole. The 2^{nd} booster can be primed before the bottom loading, during the main loading or after the main loading.

Whether the truck loads first or a bench hand primes first is a matter who is first ready. There is no base or ideal pattern of actions that a crew strives for, it is a product of who and what is available. The 2^{nd} booster is primed when a bench hand is available to prime it. If there is only one bench hand, less can be done at the same time, and the 2^{nd} booster is often primed during or after the main load, depending on how long the loading takes and how fast the bench hand can prime. If there are two bench hands at the hole, the 1^{st} and 2^{nd} booster are primed at the same time, if both bench hands arrive at the same time. There is no distinction between which bench hand does the 1^{st} or 2^{nd} booster, whomever is ready to load the booster first, becomes the 1^{st} booster. In addition, if a bench hand is interrupted or just happens to be slower during a task, the order of the actions changes without any explicit coordination. There was no norm or base pattern that people tried to create, people responded by doing what they could do with what was available.

4.7.2 The routine

While set patterns of actions could not describe the variations observed, there was some stability in how individuals, and crews, approached their work. As can already be read in the previous examples, the output of the activity, the peg with the hole depth and water level, as well as the loaded hole do repeat. These were stable for all observed variations within one crew. However, just focussing on the output alone is not enough to explain all repeating elements. For this we need to look at how people understand the problem they have to address, the structure of the solution to solve the problem, which enables creation of signs, and when shared across a crew facilitates coordination of the work.

Problem understanding

The problem understanding is what crew members believe they need to achieve and under what constraints. Problem understanding captures what a crew believes it needs to work within; the boundary of what they do not control themselves. This includes the scope of work, what equipment is available, how conditions affect their job, and what is physically possible.

Generally, the more experience crewmembers get, the more their problem understanding will converge. However, when people work in different conditions, this is necessarily not the case. When showing the same picture to members from different crews, it can evoke different ideas on what needs to happen.

One typical difference between blast crews includes the demarcation of a blast area. This shows in the different reactions from operators of different blast crews to the same picture. First, the reaction of a practitioner to a picture of a work site he had worked on that day:

"*Awh yeah, this is a good bench. They are good at that, at this site, the way they deliver them. It's all safe. It's a good work area, nothing you'd change about that. From this angle (of the picture), I'd might be worried the corner in front of the entrance was a bit tight for the truck, but I know [the team leader] has been there, and he would have done something about it if it was too tight*"

The second is an interview with two experienced practitioners from another crew:

Practitioner 1: "Is that the entrance to the shot, is it?"

Researcher: "Yes"

Practitioner 1: "So there's…. no, no, you drive straight into it."

Practitioner 2: "The chicane shit is only here. Only cause of… "

Practitioner 1: "That old SSE"

Practitioner 2: "Now he's gone"

Practitioner 1: "That's (the entrance in the picture) less work, that's good. I like their entrance"

The clearest difference is whether the way the entrance is setup in the picture is acceptable and up to standards. For the second crew there has to be a chicane, while the first crew does not have this requirement. There is also a difference in who is considered responsible for the entrance being up to standards. In the first interview the expectation is that other people in the mine make the entrance to standards, while in the second interview the crew sees it as their responsibility to build the walls and fences to their shot. This difference was confirmed in the field observations. The example also shows that problem understandings are not stable, and requirements can change over time. The problem understanding is, however, also not built from nothing for each job, and is built through previous jobs carries over to the next.

A problem understanding is not always as clear-cut as regulations that need to be complied with. Problem understanding also includes the expected conditions, tools that are available, and to what degree something can be managed. This shows people from different crews who have different ideas responding to the same picture with a muddy bench:

Practitioner 1: "It's pretty straight forward. You get messy, but there's nothing you can do about it, if you know what I mean."

Researcher: "What do you mean with that?"

Practitioner 1: "It's a weather thing, it's fairly obvious they have had rain."

*Practitioner 2: "Puddles" ****points to muddy areas******

Practitioner 1: "Yeah puddles and stuff. There's only so much of that stuff you can clean up, if you know what I mean, before it actually becomes a problem itself"

Whereas a practitioner from another crew said:

Practitioner: "I don't know how wet that is, it looks like (it's) a bit muddy, they could have cleaned that up a bit better with the loader, so it is easier with the truck to get through."

Researcher: "Wet ground is difficult to manage?"

Practitioner: "Yeah, it is a hazard for the vehicles, you get the vehicles caught, stuck. You gotta actually walk through it. You drag dip tapes through it."

Researcher: "Okay, if this was your crew, what would you be doing, or what would you be doing next."

Practitioner: "Well for starters, we would have had the rows cleaned up a lot better than that. We wouldn't have had that mud there."

The first crew sees getting messy as inevitable and muddy ground something you can clean, but only so much. Based on what they saw in the picture, they assumed it was beyond the level of what you can completely clean. The person from the second crew saw in the same picture a bench that could be cleaned with more effort.

As a general trend, we noticed that crews that worked in a mine with more rain and muddier ground, saw muddy ground as something less avoidable than crews working in mine sites in drier conditions. An indicator that their problem understanding is shaped by experience.

While interviewing practitioners from different power line depots about a single job, there were few consistent differences between practitioners from different depots, until the practitioners were asked about how an extra person or truck would change the job. The practitioners from the first crew responded that it was basically a two-truck job, and adding extra people to the job would not make a difference in the end. Practitioners from the other depot had no issue finding ways of making use of another truck and an extra crew helping them out. Through further prompts it become clear that at the first depot, crews had to return to the depot after a job was finished and wait out their time. Crews from the second depot were allowed to go home if all their jobs were done for that day. While the job set up was similar for both crews of both depots, they differed in their response to changes in resources, because they had different constraints.

Solution structure

Not only the problem understanding repeats, but also how people divide and structure the larger problem into smaller tasks, i.e., the solution structure. This includes how these tasks relate to each other, whether they need to be done in a particular order, and whether there is a particular task delegation within a crew. While the problem understanding by people across crews in similar, environments generally converges, solution structures differ more between crews, and converge within crews.

The idea of a solution structure aligns most closely to the idea of the ostensive part of an organisational routine, as it forms a recognisable repeating structure in the way a group of actors approaches a task. However, what is different about this idea is that it does not specify actions, but only interaction parts, which could be actions, but could also be in-between-states. For an experienced blast crew members measuring hole depths, the solution structure provides what the peg should look like with the hole depth and water level written on the peg. How practitioners achieve this, comes about through what is most available to the practitioner, not a matter of a set action pattern.

Crews function smoother when its members have a shared solution structure. This showed most clearly in blast crews when they had connected detonation wires, after they have all loaded the holes. As there in no going back after a blast has been initiated, blast crews check whether the wires are connected correctly multiple times after they are connected. Within a stable crew, the angles at which different wires are connected, the way spare wire is organised, and how loose ends are positioned are all uniform. Team members correct each other on deviating elements, even if they are functionally just as good, to create easily recognisable elements.

One team leader explained to us:

"There's hundreds of correct ways to tie these wires together that work. Last month we had a technical engineer here working with us. No doubt he can tie up wires, he knows more about this stuff than we do. But he put the cords different than we did, so we when we were

checking up and looked at the holes he tied up, we went like 'heh, what's this?' and you have to look carefully at how the wires run to see whether they are correct."

In this same crew, however, there were different techniques and orders of actions used by the crewmembers to connect the wires. This was nor a problem, as the different methods produced the same result. The practitioners were not even aware of the differences between the tying techniques. The in-between-states, the parts other crewmembers would act upon mattered, but the actions that produced them did not.

Within one crew, crewmembers would share solution structures, in the sense that there was an ordering of what is to be achieved first and what are acceptable in-between-states. However, there could be differences between crew members in the detail of a solution structure. For one, someone with a heavy vehicle operation license would be checking the muddiness of the ground and proceed to clean out mud with a frontloader before getting equipment to the holes, while someone without a heavy vehicle license would start by carrying detonation wires and boosters to the holes.

As a general trend, we noticed how novices tended to divide the problem into more and smaller sub tasks, showing less variation between instances. When measuring the hole depth and putting the peg into the ground, the order was almost completely stable for practitioners in their first months. Their solution structure could look like, position peg, measure hole depth, measure water level. A more experienced practitioner would immediately work towards the goal getting both the hole and water depth on the peg and could do multiple elements at the same time. By working towards a further goal, the expert could respond more opportunistically to what was available. There were also cases were experts divided a task in more set steps, for example where experts would take time to set up a job area to suit their work flow. In these cases the expert would still be focussed on goals further ahead, even if more steps would be added in.

Crews can know multiple solution structures for one task. With blast crews we noticed that there was usually one assumed structure, but with power line crews it was common that crews mentioned what method they would use, and referred to known approaches. For example, blast crews typically move between holes on a bench in different ways. The holes on a bench usually form a grid, and crews work through the holes per row. However, in some situations crews moved between holes in echelons instead.

Signs

Practitioners use their problem understanding and solution structure to see meaningful signs in an environment. This could through the relation between sign and the larger structure or conventions that arise from stability of in-between-states.

A stable solution structure leads to easily recognisable states of work. A row of uncovered holes with detonation wire coming out of them, indicates for blast crews that the holes are loaded and need to be stemmed. When all holes on a bench are stemmed and have detonation wires tied to the pegs, it means that the next step is to connect all detonation cords. This signs could be deduced from the known steps, but often will be symbols built through convention.

Differences in solution structure can explain some differences in of why practitioners from different crews interpret similar situations differently. An example showed in the practice of one blast crew that involved parking a vehicle where they expected to finish loading that day. Members of that crew would instantly give that meaning to the location of the parked vehicle. When operators from other crews saw pictures of that crew at work, they would notice and mention the vehicle, but they would not ascribe that meaning to its location. The difference in meaning ascribed to the location of the car was a matter of convention.

Besides conventions based on the solution structure, people also related what they saw to their problem understanding. Blast crew operators unfamiliar with vehicle parking practice would sometimes comment that the parked vehicle was in the way of the loading truck, showing that they take into account the way trucks move and where trucks need to be positioned to operate. Another example would be the experience people have with dealing with muddy ground and whether they see the ground as cleanable. In the cases a practitioner was asked about a picture with an activity that the practitioner had never done, the answers showed the biggest discrepancies with the actions and answers from the crew in the picture, if they could even formulate a coherent interpretation at all. Practitioners watching other teams engaged in known tasks, but with different routines or equipment could still recognise the done steps on an abstract level. This indicated that without an understanding of the problem it becomes almost impossible to understand what other people are doing.

The stability in the problem understanding and solution structure enable deviation to become a signifier. This applies for both intentional and unintentional deviation.

The case of checking up on connected wires portrayed how unintentional deviation became a signifier. As mentioned by the crew leader, when all wires are laid out the same way, he could quickly see whether they were correct and accidental deviations would stand out. When alternative options were used, he had to trace the wires and think whether that setup would work. This aligns with the idea that computational demands are traded for perceptual demands (J. M. Flach & Voorhorst, 2016). The exact shape of the deviation would not matter, but the deviation from the standard practice of the team signified crew members to the item at hand.

Teams also relied on intentional deviation to signify each other. This was not just a matter of convention, as different tokens could be used to represent the same message. Within one crew, coloured lint tied on a peg, upside down boxes, or bags tied to a peg were used to indicate that a hole was wet and needed to be dewatered. One practitioner described that it does not matter what someone uses, as long as *"you make it stick out like dog shit"*. The opposite was observed as well, where the same token was used to communicate different messages. In one crew, bags placed on a peg could indicate a wet hole and needed to be dewatered, while behind the dewatering truck, a bag on a peg could signal that crew members had started working from that point to bait holes. In these cases, the relation between the token and the message was not only based on convention, but by relating the token to the stage of the work, knowledge of what matters and what can happen during these stages of the work.

The flexibility in marking rarely caused issues. If practitioners could not infer what something meant, they knew they needed to ask fellow crew members. We did run into cases where 'leftover' equipment, stacked up unusually, left crew members uncertain as to whether it was an intentional sign. In terms of the location of the equipment, however, the practitioners deduced that if it was a sign, it was not a sign for them. During interviews practitioners would recognise intentional markers of other crews in pictures, but could usually not identify what was being communicated. If more information was given about what was going on that day or what the crew was doing, practitioners sometimes could infer the intended meaning behind the markers.

Practitioners convincingly argued that not standardising markers to that detail improved the reliability of operations. Strictly regimenting things would mean a marker could be wrong, which would create the dilemma when the right marker would not be available. If all wet holes needed to be marked with blue lint, and a practitioner unexpectedly runs into a wet hole and does not carry blue lint, the practitioner faces the dilemma of not marking the hole straight away, or using a wrong marker. A wrong marker could lead to the wrong interpretation by a fellow crew member, possibly leading to undesired actions. Obtaining the right marker could be costly, as it would take time and is not without risks. If a trip to the depot was required, this could easily cost 20 minutes in driving alone. This comes with the risks of forgetting which hole it was, being interrupted by other tasks, and other crew members interacting with the hole that needed to be marked in the meantime. Without the standardisation, crew members can grab whatever is closest at hand, making it more likely the work will be achieved as intended.

Observations confirmed that in many cases availability drove what was used as a marker. If a booster got stuck in a quarry, stones would often be placed on top of a hole, and quarries have more stones lying around than mines do, as well as smaller and easier to cover holes. In mines we have seen stones used as a marker, but exclusively when marking holes near a wall, the location where larger stones can be found on a mining bench. If a hole needed to be marked after measuring the hole depth, the hole safers (large plastic cones) were used most often, which would be laying around next to the holes at that time. Before a practitioner used a marker, they would usually scan the environment around them, also confirming the practice. While markers likely had to meet certain criteria in noticeability, suggesting that not any object could be used as a marker, there was no indication that there was a set list of possible markers.

4.7.3 Changing routines

The solution structure and problem understanding guide individual crew members within a crew, allows variation within the solution structure, and even enable adaptation in the signalling. However, as routine can change, the solution structures and problem understandings can change as well.

Change can happen quickly if all agents involved can understand the changed solution structure or problem understanding. This does not necessarily require explicit coordination. In one job, a blast crew in mining had to work on an unusual small rocky bench for them and needed to put the stemming in the holes without the specialised stemming truck they usually used. The crew started out with using shovels to carry the stemming across the bench. This was a frustrating process as stemming often fell off the shovel in the process. Some crew members started to swipe the stemming towards the hole with their boots, but this was not very efficient either. One crew member arrived later on the bench, as he came from another job. This crew member picked up an empty nylon wrapping bag, lying near the entrance of the bench, and started to fill this bag with stemming, then walked the bag over to a hole, and in turn emptied the stemming from the bag into the hole. As other crew members saw him do this, they started to use other leftover bags. As one person had just filled his bag, another crew member was waiting next to him with an empty bag on the ground.

The person with the shovel started to fill up the other crew member's bag. From this, a task distribution arose where two people stood with the shovels near the piles of stemming filling up empty bags, while the other crew members walked back and forth between the stemming piles and holes with the bags. This all happened without any verbal communication to coordinate this change. Later interviews uncovered that the crew had used those wrapping bags before to carry stuff in, but had never used the bags specifically for stemming. During the work, before the researcher inquired about it, the initiating crew member joked to the others that they stole his idea and should pay him for it, confirming that the other crew members had copied from each other, rather than it being a known pattern or something that was instructed. The development of a new routine requires that crew members understand the adjusted solution structure and problem understanding, not necessarily explicit coordination.

For a routine to change, the solution structure needs to change. This means a change in a path to a (sub)goal within a routine, is not a break from the routine, as it fits within the freedom the solution structure provides. In the shovel and bags example, as long as one operator did the shovelling and carried their own bag, this does not break the shared solution structure of the team. However, when one person starts shovelling bags constantly, while another person walks bags back and forth, this is a change, as the task is now divided and structured differently into smaller tasks. A change in action only becomes a change in routine, when it affects other steps as well; when one changed element reverberates to other parts of the routine.

As change of a routine equals reverberation, we have noticed that all routines changes come with a process of searching for a fit, followed by a period of stabilisation and standardisation.

The process of searching for a fit involves finding a new solution structure that leads to smooth operations. During this process, it is more common to observe variations in the in-between-states than stabilised routines. This search is affected by what goals and constraints crews have, like regulations and work conditions. As general reasons why practices between crews differ, individual preferences were often mentioned. However, when discussing specific examples of differences in practices, crew members would almost never mention individual preferences and usually mentioned conditional factors or features of the team, even if this information was not visible in the pictures. The factors and features mentioned held up in all cases when triangulated with other data. In our data, the 'fit' with the goals and constraints played a bigger role in what made people consider a routine 'good' or 'better', than individual preferences.

This is not to say that individuals have no influence on the development or search for new routines. When crews faced unfamiliar conditions, individuals would often bring up things they had learned when working in other crews, which the crew would then adopt. In line with this, crews did not always respond the same way to changes. Across one region, a procedure was rolled out to blast crews around how to tie up loose wire between a loaded hole and the peg, to reduce tripping hazards. While different crews received the same instruction, crews did not end with the same 'knot', nor did they add the tying in at the same moment in their routine. In some crews, the crew members immediately tied up cords after the hole had been loaded, in other crews, they tied the wires together after they had checked the electric functioning of the wire and clip. In the former crews, the tied wires were an indicator the hole had been loaded. In the latter crews, the tied wires became a sign that the functioning of the wires had been checked. The structure of the sub tasks, the in-between-states, and the sign meanings could change as crews search for a fit as a routine adjusts.

During the stabilisation of a routine, standardisation of signs and in-between-states was observed, even if this was not necessary for the teams functioning. Knots and tying methods would converge, even though functionally multiple options were possible. In quarries, blast crews have to communicate hole numbers consisting of a letter and a number over large distance in a noisy environment. For this, they have developed a sign language, where they point to a body part that starts with the same letter as the hole number followed by holding up the fingers for the number. So, for A2, they would point to their arm, followed by holding up two fingers. The body parts used for each letter were standardised within crews, but some varied between crews. Because of the underlying rule, operators had no trouble reading signs from other crews. As the researcher inquired about the sign language, crew members themselves became curious and started to experiment with new signs. Practitioners started to try never before used signs, following the same rule. Their fellow crew members had no issue interpreting these newly invented signs. While the familiarity with the task in general and location of the hole likely helped to interpret the new signs, it confirms that standardisation is not required for teams to function, yet it continuously happens, suggesting a drive for simplicity and optimisation.

4.8 Discussion

4.8.1 Summary of a routine

From our results, we can craft a novel explanation of how stable crews perform routines and how routines shape the performance of crews. This explanation incorporates individual decision-making elements, larger repeating patterns on the team level, and how the two influence each other.

On an individual level, people develop a model of what their task is and what they need to achieve—their problem understanding. They divide and structure the problem into smaller tasks, forming a solution structure. The solution structure provides the goals people are oriented towards during their work. Solution structure maps directly to Feldman and Pentland's (2003) 'ostensive aspect', but specifies what the ostensive does and does not capture. The solution structure provides goals and in-between-states.

As in Feldman and Pentland's (2003) theory, the solution structure guides the performative aspect of a routine including improvising details that are appropriate to the context. The solution structure does not provide the actions, nor is does it divide the problem into an infinite level of detail.

Below the level of detail contained in the solution structure, people respond opportunistically to what is available to them. As a general trend, as people become more experienced, people can reduce the specificity of the solution structure, giving them more freedom to adapt within the structure.

The combination of the problem understanding and a solution structure allows people to give meaning to their environment. This can be through relating signs to their problem understanding or solution structure, but can also lead to direct associations between tokens and a message. By having a stable problem understanding and solution structure, people can develop norms and expectations about their environment. Deviations from these norms, in turn, are easier to identify.

Repeating patterns in teams arise because team members hold similar problem understanding and solutions structures. The overlap gives rise to the signs that coordinate individual team members. When all crew members work towards the same larger task and divide that task into the same smaller steps, each member can determine how to contribute. As long as an individual's problem understanding is appropriate, and their solution structure fitting for the situation, all the team's actions should combine into a completed task.

Problem understandings and solution structures do not need to be exact replicas across different team members. The solution structures of different team members need to match at the points where the work of crew members interacts. These points of interaction are often in-between-states, where one crew member picks up the work of another. However, they can also involve direct interaction (such as hand signals). Between these points of interaction, crew members have freedom to adapt. So long as crew members apply similar solution structures, differences in problem understanding remain latent. However, when a solution structure does not work to complete the task, and must be updated, crew members with different problem understanding may have different expectations and decide on different courses of action.

Crews can handle (some) diversity in solution structures between members, but have a tendency towards similarity. One mechanism behind this tendency is that crew members correct other crew members if their actions reflect a different solution structure. At the team level, it is beneficial to have shared solution structure as it creates shared expectations. With shared expectations, the same elements will stand out as deviations to anyone in the crew. This creates the possibility to create intentional deviation as a signifier—for example the use of upside-down boxes as hazard markers.

Routines can change, because problem understandings and solution structures can change. The performative aspect is always improvised, but these improvisations can adjust parts of the solution structure. Experience in successfully or unsuccessfully performing a solution structure can evolve problem understanding. Change does not require explicit coordination but does come with a process of searching for fit and stabilisation of a new routine. Individuals can only unilaterally update their solution structure if it still provides consistent interactions with the rest of the team.

4.8.2 Conclusion

This study offers an in-depth view of how routines play out across a team, explaining how the path-dependency comes about in a team. We discuss shortly how our explanation aligns with theories on individual decision making, team processes, and organisational routines literature and how it contributes to or contradicts these. In addition, our explanation raises doubts about the distinction between heedful interrelating and routine action, to which we provide an alternative view.

Individual decision making

Our explanation aligns with individual cognitive literature on that people are goal focussed (Rasmussen, 1983), with goals structured within each other (Powers, 1973), and use mental models to give meaning to what they observe (Hutchins, 1984). As an individual perspective within teams, it has been mentioned that schema driven decisions make crew members predictable to each other, which aids in coordination (G. A. Klein, Feltovich, et al., 2005; G. A. Klein et al., 1993). What our view adds is how team level processes shape goals and goal hierarchies of individuals, and how schema driven decisions not only help coordination by interpretation of others but also help individuals generate actions that support completion of the combined task.

Organisational routines

Our view aligns with the organisational routine literature in that there is stability on how multi-agents groups perform a task (March & Simon, 1958), that routines form effortful accomplishments that aid in coordination and communication (Pentland & Rueter, 1994), and that variation can be absorbed by routines (Turner & Rindova, 2012). Our study provides an answer to the question of how different actions link together as a routine is performed (Pentland & Hærem, 2015). In addition, we give an explanation for how experts can become more adaptive and flexible in their working method, rather than more rigid, as the idea of imprinting an ideal image suggests (Turner & Fern, 2012), and how teams coalesce that have been working together (Turner & Rindova, 2012).

Our results also contradict theories in the organisational routines literature on some points. Our explanation says that beyond the solution structure (that consists of goals and in-between-states) there is no base or ideal pattern of actions that guides practitioners as they perform their work. Actions come about opportunistically, through what is most available in the environment, or what the environment affords. This challenges the definition of routines as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (Feldman & Pentland 2003, p.95), at least when this definition is applied to the ostensive part of an organisational routine. From this it also follows that organisational routines cannot be infinitely decomposed to smaller levels of analysis as claimed by Pentland and Hærem (2015).

Heedful interrelating

Our study challenges the idea that routines are different from heedful actions; relating current actions to the larger task is inherent to how routines are performed. In contrast, stability of a solution structure in fact helps people be more sensitive to current operations.

However, our observations support Weick and Roberts' (1993) suggestion that rotating team members can increase changes in practices and that certain moments are more shaped by a search for actions that work for the problem at hand. Therefore, rather than heedful interrelating versus routine interaction, a more fitting distinction would be between robustness and fluidity of a routine. Different team features and conditions can make a routine more robust or more fluid. When practitioners in a team have a shared problem understanding, a shared solution structure with minimal shared points, and redundancy in possible signs per interaction, a routine is robust. The routine can handle and show variation, without having to break the solution structure guiding individuals. The ability to handle a wide variety of situations also means it is unlikely that a team will be prompted to search for new solution structures, even when the current setup is sub-optimal. A lack of diversity between crew members also makes it less likely that crew members will come up with different solutions. This makes a routine robust, in that it is both capable of handling variation and unlikely to change.

Teams where the members have diverging problem understandings and solution structures will require more explicit coordination. At the same time, they are more likely to search for the best fit given the current circumstances and larger goals. This makes the routine change more often and more likely that the agents involved will consider more possibilities, making the routine more fluid.

Both robustness and fluidity can be desirable when trying to engineer resilience into a system (Hollnagel et al., 2006), depending on the circumstances. From this it follows that stable crews are better suited for jobs with known variation, while changing crew membership can be beneficial in situations with variation not encountered before, or when searches for more optimal solutions are desired. This aligns with the idea that team diversity and team member rotation increases problem-solving capabilities (Gorman et al., 2010; Hong & Page, 2004) and that common ground aids in coordination (G. A. Klein, Feltovich, et al., 2005), but more research is needed to verify this view.

4.8.3 Implications for research

Beyond adding to the theory, the concept of a solution structure facilitates the study of changing routines and long-term adaptation in socio-technical systems. The concept of solution structure allows distinguishing between variation within a routine and changes of a routine. It enables separating out noise when studying how routines change over time. The difference between variation within a routine and change of a routine has been raised for organisational routines (Cohen, 2007), but is also relevant for research on ideas such as (procedural) drift (Dekker, 2011; Snook, 2000) and normalisation of deviance (Vaughan, 1996), within socio-technical systems.

4.8.4 Limitations

Our explanation has been verified with blast crews and power line crews. We assume that the basic process applies to stable teams in general. However, there are likely to be differences across industries in how and where standardisation takes place. In my study standardisation played out within teams and specialisation happened between teams, but this is not necessarily universal. For pilots, for example, with less stable teams, standardisation might happen on an organisational or even cross-organisational level, as suggested by the ideas of airmanship and airlineship (Haavik et al., 2017). In addition, our study involved mostly homogeneous crews with limited specialisation. In crews where members have different specialisations, it is possible that shared solution structure and problem understanding are not strictly necessary. It is not necessarily similarity that creates compatibility.

4.9 Outcome for the research aims

Similar to the early HRO studies, by studying a successful group, a process of organising work was found. My explanation mostly focuses on the coordinative aspect, but can also be used to explain how successful outcomes are reached in complex environments. It can even be linked to HRO theories, in that routines can assist in creating sensitivity to operations via the idea of deviation. In turn, this can increase mindfulness. Like, HRO, my explanation acknowledges that people's actions change in relation to the environment, when it is not required. Unlike HRO, however, I stress that repetition on some levels is useful and beneficial. In that regard, I have found a new process through which to organise for safety.

In addition, my findings provide suggestions on how to organise blast crews. Firstly, within blast crews, it provides an explanation blast crews can use, however, as this appears to be how crews self-organise, teaching it will likely be of limited value. Other elements, however, can be adjusted to suit this style of organising, such as team stability and specification levels of procedures.

Chapter 5: **Turning repetition into expertise for resilience**

5.1 Statement of contribution to co-authored accepted paper

This chapter includes a co-authored paper. The copyright has been transferred to the Oxford University Press, however, authors have the right to include their article in a thesis or dissertation that is not to be published commercially. The bibliographic details of the co-authored paper, including all authors, are:

Havinga, J., Bergström, J., Dekker, S., Rae, A. (in press) In Ward, P., Schraagen, J. M., Gore, J., Roth, E. (Eds.) The Oxford Handbook of Expertise: Research and Application. *Oxford, Oxford University Press*

My contribution to the paper involved:

The creation of the outline for the entire paper; The incorporation of feedback on outline from Dr Bergström, Prof Dekker, and Dr Rae; The review of the literature; The writing of the first draft, excluding the resilience section; The incorporation and alignment of the resilience section into the rest of the paper; The editing of the first draft based on editor feedback; The editing of the manuscript based on feedback from Dr Rae, Dr Bergström, and Prof Dekker.

(Signed) ______________________________ (Date)____________

Joseph Paulus Havinga (Corresponding author)

(Countersigned) _________________________ (Date)____________

Dr Andrew Rae (Principal supervisor and co-author)

(Countersigned) _________________________ (Date)____________

Prof Sidney Dekker (Principal supervisor and co-author)

5.2 The rationale for the study

The fieldwork showed that blast crews have routines where they repeat a structure used previously, but routines cannot only be explained in terms of repetition or convention. Any observed repetition happened on a background of expertise with the task in general, which showed most clearly when teams needed to adapt their practices. In some ways, the repeating elements in the solution structure provide the glue that binds individual team members together, however, the stability relies on the expertise in the how to adapt to the environment, as well as, what can and cannot be adapted to keep the team working. This makes expertise a key element for organising blast crews to be resilient, which leads to the larger question of how expertise relates to resilience and how expertise can be managed for resilience.

Expertise and Resilience

5.3 Abstract

Resilience Engineering (RE) has changed the value of expertise from meeting required standard, to how it helps organisations to adapt. This chapter discusses the origin of the concept of resilience and how it has it has been applied to socio-technical systems within the safety domain. From there, we review the current literature to explore how to manage expertise, considering both its possible good and bad effects, to engineer resilience into organisations. For this, we study how this applies separately to frontline workers, teams and management, and on a systems level.

5.4 Introduction

Safety has a long history of reducing risk by limiting variability. In that approach, expertise was about meeting the standards set in the design phase and having the ability to keep a system working as intended. Resilience Engineering, a new approach to safety has a different view on the benefits of expertise. Resilience Engineering values expertise as it can help adjust operation to varying conditions. This builds on the law of requisite variety, which says that to deal with varying conditions, a system must be capable of varying responses (Ashby, 1957). There are still desired outcomes, but experts are not necessarily expected do the work as was planned, and can adapt to the circumstances.

In this chapter, expertise is defined as the change in how an agent approaches a problem or job based on experience. An "agent" is not limited to humans, and can also include people combined with supporting technology, a team, or larger organisation. Experience is considered everything an agent goes through, which includes time on the job, training, and even conversations had about work. Expertise opposes built-in or designed ways of functioning, and innate capacities, such as talent. Expertise is generally linked to improved performance and more sensitivity and adjustment to conditions, things desired for resilience. However, these benefits are not integral to the definition, in fact, in this chapter we will consider the evidence for and against expertise providing such benefits. We will also consider the current state of knowledge about how expertise can be best managed to engineer resilience into different levels of an organisation.

Before we go into this, we first explain what resilience is, its history, and how it is applied as alternative to safety.

5.5 What is resilience

The notion of resilience is a multifaceted one. “Resilience” is used as a metaphor, a theory, a set of capacities or even a strategy (Norris, Stevens, Pfefferbaum, Wyche, & Pfefferbaum, 2008). The common denominator seems to be that resilience is always related to some kind of coping with stress (Kolar, 2011).

While the term resilience originated in material physics as a material’s ability to regain its original shape after having been put to stress, resilience as used in contemporary (social and organisational) science comes via psychology; seeing resilience as a socio-psychological agent’s ability to thrive despite adversity (Garmezy, 1971; Garmezy & Streitman, 1974). With this meaning resilience has been a target of psychology and health research since the 1950s, typically using war traumatised children or veterans as targets of analysis (Tyhurst, 1951). The central question to answer in such resilience research is “why do some individuals seem to thrive despite the great adversity they have been experiencing?” In the 1970s resilience was introduced as a concept also in ecology; which was defined as an ecosystem’s ability to absorb and adapt to stress without going into a qualitatively different system state (Holling, 1973; B. Walker, Holling, Carpenter, & Kinzig, 2004). Later, the ecological resilience research included socio-ecological systems; expanding the system boundaries to include research on human interaction with ecological environments (Gunderson, 2010; Gunderson & Holling, 2002).

Following the 9-11 terror attacks on the US, resilience became a target of public policy. Citizens, communities, cities and entire nations are now supposed to be resilient in the sense that they should have the preparedness to adapt to any unforeseen natural or antagonist event (J. Walker & Cooper, 2011).

Looking at the relatively long history of resilience research and policy the safety science community was rather late to adopt the notion of resilience in the early 2000s. Resilience Engineering was introduced in Woods' (2003) testimony on the future of the National Aeronautics and Space Administration (NASA) to the U.S. Senate Committee on Commerce, Science and Transportation, a committee led by Senator John McCain. Woods introduced it as a school of thought dedicated to "help organizations maintain high safety despite production pressure" (p. 2).

In the book following the first Resilience Engineering symposium, Resilience Engineering was presented as a dedication to improve and engineer more resilience into systems. This was born from frustration with safety thinking based on counting errors and avoiding negative events. The new analytical path focussed on how organisations anticipate and adapt to the complex and dynamic risk landscape (Hollnagel et al., 2006).

To the school of Resilience Engineering; resilience is something positive and desirable, representing "the optimist stance and its agenda is to develop ways to control or manage a system's adaptive capacities based on empirical evidence" (Woods & Branlat, 2011, p.128). With this, Resilience Engineering stands against more negative safety schools such as Normal Accident Theory (Perrow, 1984) or Normalisation of Deviance (Vaughan, 1996), which suggest that complex systems are untameable and there is nothing substantial that can be done to prevent accidents in these systems. Instead, with Resilience Engineering, scholars' attention is guided to finding positive adaptive capacities, coping strategies and trade-offs, things that often coincide with expertise, that make systems function despite pressures, goal conflicts and dynamic risks; rather than the traditional focus on safety as the absence of negatives, such as errors or violations.

A question that precedes any analysis or engineering of resilience is what the subject of resilience is, as in, who or what has resilience or creates resilience with their actions. For material science and psychology, the subject comes with the perspective of their science. (Eco)system scholars, on the other hand, need to draw system boundaries before a subject, becomes an entity possible to study. For example, when studying resilience, do you look for it in a in a tree, or a forest, and is it useful to consider migrating birds as part of the forest? The interactions and relations between system levels often become more fundamental than the resilience of individual levels (Bergström & Dekker, 2014).

Resilience Engineering tends to focus on sociotechnical systems; systems that consist of people and technology interacting and that, to some degree, self-organise. However, there are various levels in which scholars target their analysis of resilience within sociotechnical systems. These can range from the resilience of a high-risk organisation as a whole, to the (more typically targeted) resilient behaviour of individual system actors (Bergström et al., 2015). At the level of frontline actors, experts are often studied in how they respond to incidents. This includes critical incident control room operators coping with adverse events in terms of flexibility (Grote, Weichbrodt, Günter, Zala-Mezö, & Künzle, 2009; Owen, Healey, & Benn, 2013) and adaptive strategies (Gomes, Woods, Carvalho, Huber, & Borges, 2009). A somewhat higher level of resilience analysis is the research focusing on the resilience of teams and management (e.g. Furniss, Curzon, Blandford, & Furniss, 2016; Gomes, Borges, Huber, & Carvalho, 2014; Lundberg & Rankin, 2014; Rankin, Dahlback, & Lundberg, 2013; Tveiten, Albrechtsen, Wærø, & Marit, 2012). The even broader level, and the original analytical aim of Resilience Engineering, is the organisational level.

Resilience needs to be studied at different levels and across different levels, as resilient behaviour at one level can decrease the resilience at other levels, which builds into ethical dilemmas. Many organisations constantly rely on their staff's expertise to adapt beyond what a system was designed to handle, but without formally recognising or appreciating such effort. This hides potential learning opportunities, as the adaptation can make the stress ignorable for others. To give one example; a Scandinavian hospital emergency ward was experiencing a high load of incoming patients[3]. Being a typical (and seemingly increasing) problem of Scandinavian healthcare, the staff coped by placing the incoming patients in beds located in the corridors. At one point there simply were no more alarms to hand out to patients for them to sound in case of emergency. Having never experienced this before, the nurses came up with the solution to start handing out cutlery and pots to the patients for them to have something to make noise with in case they needed to attract the attention of the staff. Indeed, this might seem like a great case of resilience, i.e. experts adapting to ensure the functioning of the system.

The problem, however, is that by relying on the adaptive capacities and behaviours of sharp-end experts; there is a risk that the system learns the wrong lesson. To use the example above; highly resilient experts (nurses) might use their adaptive strategies and capacities (invent an alarm system of cutlery and pots) to hide the stress (the ward kept working and delivered care) that is put on the system (the overloaded ward) and contribute to the preservation of system vulnerabilities (an acceptance of an overloaded emergency ward) and further drift towards the boundaries of possible operational performance (a collapse of emergency care at the hospital). This in turn can prevent long-term investment of management to increase the capacity of the system, making the organisation as a whole less adaptive.

A more ethically sound way to embrace resilience, would be to raise the question of whether (and how) the system provides the adaptive resources and capacities needed for the frontline experts to adapt to the situations and challenges they encounter in their daily lives. This directly creates care for the experts who are placed to guarantee the system functioning and reinforces the connection between the experts and the higher organisational levels. As a product, accountability is then shared by the entire organisation, not just the people that happen to be closest to the action.

---

[3] The ultimate state of this system became so weird that it was reported in the media (which is the reason that we at all know about it) and was later followed-up by the regulatory agency.

We will look at how to manage expertise to enhance resilience on the level of frontline workers, teams and management, and the level of systems. First on the level itself and where known, we make links to other levels.

5.6 Frontline workers

Sociotechnical systems are never fully prespecified or planned. This means people at the front line will always have to make decisions. Managers plan and design work according to how they imagine that work is carried out, which will always be different from how work is actually done. The imagined work is never as complex and detailed as the real work. There will always be mismatches and unforeseen variations (Hollnagel, 2012). This is captured in the gap between work-as-imagined versus work-as-done (Dekker, 2005b) and builds on older concepts from francophone ergonomics, Tache and Activitée (Hollnagel, 2014). It is up to frontline workers to overcome these gaps and holes in the planned work. In that regard, frontline workers inherit bad systems and form the last line of defence against accidents (Reason, 1990). To deal with the gap, humans use heuristics, work-arounds, and understanding of the system, which reflect a worker's expertise. This means any socio-technical system relies on expertise of frontline workers (Hollnagel, 2014). In the following sections, we first discuss whether expertise creates resilience and then discuss how training, procedures, and automation can be managed to enhance resilience through expertise.

5.6.1 The effect of individual expertise on resilience

In general, expertise is related to improved performance, which is desirable for any system (Farrington-Darby & Wilson, 2006). The question whether expertise always leads to more adaptiveness, that is, more varied responses to changes in conditions, is not as clear-cut. We explore this from the perspective of attention to the task and fixation of strategies or ideas of frontline workers.

In a variety of research, it has been shown that with expertise, people develop attentional abilities (Farrington-Darby & Wilson, 2006). At the same time, it has been said that with experience people start paying less attention to their job, in the idea of complacency. Complacency as applied in safety research, does not have a single definition, but is often described as a state of low suspicion and trust that things will go right, which builds during the use of a system (Moray & Inagaki, 2000). That operators in a simple signal task, such as radar, decrease in vigilance after they become familiar with a system has long been established (Mackworth, 1948). It has also been found that in other more complex simulated tasks that as people become more familiar with a system, they start to trust it more, and this predicts how closely they monitor that system. If a system shows faults, trust declines, and monitoring increases again (J. D. Lee & Moray, 1992; Muir & Moray, 1996). However, while monitoring frequency decreases, performance has not been found to decrease in these more complex tasks (Moray & Haudegond, 1998). This suggests that it in complex tasks the decrease in monitoring reflects a more efficient strategy, rather than a worse one. Trust functions as the mechanism that calibrates monitoring frequency (Moray, 2003) as expertise builds.

There are many concepts around the idea that people can get stuck on certain ideas or courses of action because of previous experience, including the Einstellung effect (Luchins, 1942), functional fixedness (Duncker & Lees, 1945), cognitive tunnel vision (Cook & McDonald, 1988), cognitive hysteresis (Norman, 1986), and fixation errors (de Keyser & Woods, 1990). While there are differences among these concepts, one reason for the variety of concepts for similar ideas is that they originated in different types of research. The Einstellung effect, for example, was framed after experimental studies, where participants were given a task of filling water jars. They found that if participants were first given tasks where a more complicated strategy was required, participants would stick to this approach, even when new tasks allowed simpler, more straight-forward strategies. Participants that were not given the more complex tasks first, used simpler strategies (Luchins, 1942). Fixation errors, on the other hand, was framed around returning patterns identified in accident investigations and more naturalistic studies of work, where people stuck to one course of action or diagnosis, even when information came in that this was not fitting (anymore) (de Keyser & Woods, 1990). Fixation errors are not necessarily caused by expertise, but it was common that people relied on their expertise to explain away signs (G. A. Klein, Pliske, Crandall, & Woods, 2005), suggesting that the expertise can enable fixation. Interestingly, it has also been found that expertise can help break fixation. In a comparison of doctors with varying levels of expertise on diagnostic tasks, it was found that the most experienced doctors were better at recognising that their initial diagnosis did not fit as new evidence came in. In-depth analysis of these results suggested that this is because the more expert doctors have richer sets of known responses and are better at recognising when things are not 'standard', thus, recognise the need to look for alternatives ideas (Feltovich, Spiro, & Coulson, 1997; Spiro, Coulson, Feltovich, & Anderson, 1988). This is in line with more recent research, where it is was found some experience can make people more rigid, but a lot of experience makes people more flexible (Bilalić, McLeod, & Gobet, 2008).

For both the attention and fixation, it is unlikely that possible downside of expertise outweighs the benefits for resilience. While for single cases, there is the possibility for expertise to have negative effects, as a general rule, investing in expertise benefits the resilience of frontline workers.

5.6.2 Training

Arguably every successful training program increases resilience of individuals in some way, either by giving people new ways to deal with something familiar, to deal with a higher task load, or to handle new situations. There is no question whether training can increase resilience in specified areas. In this section, we explore whether training can create an expertise for being resilient in general. Therefore, we will focus on training that allows people dealing with unknown, unforeseen situations, and going beyond what was specified during the design of a system. Training for this purpose can aim for skills that are useful in unexpected situations, training in dealing with uncertainty, or training to teach individuals to continuously reflect on and adjust their work practices.

There are multiple types of training that prepare individuals to deal with the unknown or aim to teach skills applicable to many situations, with the most popular being Crew Resource Management (CRM) training. CRM originated in the cockpit, but has been applied to many industries. The development of CRM started in the late 70's (Kanki et al., 2010; Maurino & Murray, 2009). CRM is older than Resilience Engineering, which shows in that CRM is often promoted as a remedy for human error, while Resilience Engineering was created to move away from the term human error. Many CRM programs aim to train individuals skills, such as decision making, situation awareness, cooperation, leadership, which are applicable in many task and workplaces in both routine and emergency situations (Flin et al., 2008). The idea of making people more capable for both normal work and unexpected situations aligns with the goal of Resilience Engineering.

While CRM training is popular and has been regarded as a major success, there are some critical notes that can be made about it. CRM has not always been defined clearly and training programs with different goals and philosophies are grouped under the label CRM (Salas & Burke, 2001). Because of this, claims of the successfulness of CRM in general do not necessarily reflect to single training programs. The evaluative studies that did find effects have mostly found effects on attitudes and few on changes in behaviour or effectiveness of performance (Salas, Wilson, Burke, & Wightman, 2006). In addition, many of the evaluation studies suffer from methodological issues (Edkins, 2002), including no comparison groups, not ruling out random variation, and not analysing questionnaire data according to their designed theoretical constructs. The uncertainty around what CRM 'is' and its effects raise doubt as to whether it teaches the general skills some claim it does. To gain more answers of whether CRM training does train widely applicable skills, research needs to test whether the trainees acquire these skills, whether they are applied consistently in both trained and non non-trained situations, and whether the use of these skills leads to better performance.

Besides training aiming for certain identified skills, there are also activities frontline groups can conduct to enhance their expertise and resilience, such as deliberate performance (Fadde & Klein, 2010), and collaborative cross-checking (Patterson, Woods, Cook, & Render, 2007). These activities allow frontline workers to reflect on their own or each other's practice.

There are no studies on whether these practices lead to fewer accidents or increased performance in typical safety-critical industries Resilience Engineering is applied to. Their introducing papers (Fadde & Klein, 2010; Patterson et al., 2007) do show that these processes change the behaviour of people and draw long-term lessons from it. That people change their behaviour based on their situation and goals, suggests that it creates resilience. The unanswered question is whether the individual changes are an increase of the resilience of the larger system, or a trade-off between resilience in one area for resilience in another area. With these practices the aim is on specific processes and adaption of those, which is also expected to lead to the development of worker expertise applicable to other situations. CRM, on the other hand, aims to instil a wider knowledge base, which practitioners will lead to change in behaviour in multiple situations. For both practices, there is not enough evidence to support that these works as suggested, nor that it creates an expertise in being resilient.

5.6.3 Procedures

Besides training, there are work conditions that affect resilience through expertise. Procedures, and how strictly workers have to comply with them, have a major impact on expertise and subsequently resilience.

Procedures and expertise have a complicated relationship. Procedures can capture strategies of experts and get other people to perform better. Scientific management was largely based on this idea (Taylor, 1911). In safety, rules and procedures are often made to comply with (as opposed to plan, coordinate, or prompt actions) and seen as a barrier to stop people from making poor decisions that have been found in hindsight of accidents (Hale & Swuste, 1998). In this way, procedures can be seen as a protection against a lack of expertise. However, procedures also constrain the possible benefits of expertise. Procedures tell in a mostly context-free way what a person should do. For novices, this can be helpful, but experts are generally more sensitive to context and pick up more relevant features of a situation to adjust to than a procedure specifies (Dreyfus & Dreyfus, 1980). On top of that, experts build an understanding, or mental model, of how things work. They can use this understanding to determine what to do, even when they have never encountered a situation before. To illustrate this, Patrick and Haines (1988) compared trainees from two different training programs on their performance at fault finding in a chemical plant simulator. In one training program, the trainees were taught action procedures on what to do in what situations, whereas in the other training program, the trainees received an explanation of how technical system worked. The people trained with procedures were faster at dealing with situations matching the procedures, but the people who received the explanation of the underlying system were better at dealing with novel cases. This does not only matter for training, but also how expertise develops on the job. Organisations enforcing compliance with procedures can be harmful to the development of an understanding of the job. To apply procedures, a worker does not need a deep understanding of the job, nor is the worker prompted to develop this deeper understanding. Exploring opportunities to understand the job may even get workers in trouble, as experimentation means varying from standard procedure, which would be flagged as non-compliance during an audit, with possible negative consequences for the worker's career (Reason, Parker, & Lawton, 1998). By a strict focus on compliance, the experiences of workers become less diverse, limiting the development of expertise.

5.6.4 Automation

Besides procedures, worker's resilience and expertise is affected by the technology around them, especially automation. While attempts to increase expertise usually relates to a bigger reliance on human decision making, i.e. the rationale behind the introduction of automation is often the opposite, to keep 'unreliable' humans out of the system. Similar to procedures, automation is specified towards certain design limitations. Automation does not think of different responses for new situations or decide to reorganise itself over time, like human experts can. While automation might increase the capacity of what a system can handle to understood stress, it decreases the ability of the system to go beyond what it was specified to handle.

The problem for resilience, however, goes further than the procedure-like decision making process of automation; the automation has a larger effect on the resilience of a system. Operators that first did the job now become supervisors of the automation. Where workers before were involved in every step of the process, aspects of the process can now be hidden from the new supervisors. Feedback from the automation is often minimal and delayed, making it hard to monitor the system, predict what it is going to do, and build general understanding of the task. This makes it hard to develop effective expertise for the new job (Kahneman & Klein, 2009). On top of that, the worker's skill to manually take over will erode if not practised somehow (Bainbridget, 1983; Sarter, Woods, & Billing, 1997).

This can make the larger system vulnerable to situations that automation cannot handle. When part of the automation system fails or the automation is compensating for something, like an autopilot trying to keep an airplane steady when there is an imbalance, it does not necessarily inform the human supervisor of this. When the automation suddenly cannot handle it anymore and hands control back to the supervisor, the supervisor is given little notice and is immediately thrown into a critical situation where it might be impossible to establish control (Sarter et al., 1997). The automation can make the larger system less resilient by taking the expertise away from the human and hindering the development of expertise for the new supervising task.

This, however, does not mean that there is no place for automation in resilient systems. The solution to the brittleness of automation is in how it interacts with people. People can be trained in how to deal with automation, focusing on understanding how automation works, opposed to just input and output knowledge. In addition, the automation needs to be designed as a team player, in that the automation is predictable and directable by signalling what it is doing. The automation needs to inform the supervising humans when it is stretched to near its capacity (Bainbridget, 1983; G. A. Klein, Woods, Bradshaw, Hoffman, & Feltovich, 2004; Sarter et al., 1997). Overall, given enough feedback and transparency supervisors can develop expertise in handling the automation. This does not only apply for automation, but for all technology introduced to people. It is not only a question of how well people perform with technology, but as conditions and goals can change in adaptive systems, people also need to be able to quickly learn how to use technology in new ways and recognise when a system is performing near its boundary (Hoffman, Marx, Amin, & McDermott, 2010). Rather than just focussing on maximum performance, for engineering resilience, it is important how quickly the system, including human operators, can readjust to new goals and conditions.

5.6.5 Team and Management

There is a large body of literature on High Reliability Organisations (HRO) that is highly applicable to how team and management practices influence resilience. The HRO studies can be traced back to a group of researchers at the University of California, who started to study how organisations that have to deal with high risk manage to achieve safe and reliable outcomes (La Porte & Consolini, 1991). This included studies on flight deck carriers, nuclear power generation, and air traffic control, to identify what the 'good ones' do (Weick et al., 1999). In a sense, they aimed to capture what the expert organisations do to deal with risk.

Originally, the HRO work was descriptive, but later the work turned more normative and started to provide advice on how organisations should organise for reliability (Weick et al., 1999). This idea assumes that there is something different about those organisations. Resilience Engineering takes a different approach on this and does not think success and failure have to come from different processes and instead focuses on understanding work (Haavik et al., 2016; Le Coze, 2015).

However, most advice of HRO does align with the goal of engineering resilience into sociotechnical systems. Dekker and Woods (2010) have even described Resilience Engineering as the action agenda for HRO. In line with that suggestion, we will discuss how managing expertise matters for resilience, along the principles of HRO that most strongly relate to expertise. These principles are deference to expertise, pre-occupation with failure, and commitment to resilience. We end this section with an exploration of whether expertise can be detrimental to resilience on the team level.

5.6.6 Deference to expertise

Deference to expertise means that the decision authority is given to people who know the most about the operations and the local situation, opposed to hierarchical position in the organisation or standard procedures (Weick & Sutcliffe, 2007). Rochlin, La Porte & Roberts (1987) described how on aircraft carriers, even the lowest ranking personnel could suspend an operation if they believed it was unsafe. This closely aligns with Resilience Engineering idea of changing from original plans, adapting to circumstances, and facilitating people to do so. A lack of listening to experts has also been linked to development of accidents (Dekker, 2014a). However, the wide support for deference to expertise does not mean deference to expertise is easily put into practice. For one, 'deference to expertise' raises the question where do you find the expertise to turn to for a decision? A designer of a piece of machinery may have very detailed knowledge about the inner-working of a machine, but maintenance technicians may have a better understanding of specific problem or how a machine is used in practice. Even if there is no doubt on whom the best suited expert is, the problem can still be beyond the expert's expertise. In that case, should that expert still be given the full decision authority?

Besides who to turn to and how much decision authority should be given, there is another dilemma created for managers by deferring to expertise. Giving up control and formally recognising people as the experts can come with undesirable consequences. Experts are recognised by organisations for their specialised technical competence, assertiveness, self-confidence, and ability to go beyond standard responses. However, these characteristics can also leave the experts insensitive to the larger organisational goals (Girard, 2005), a feeling they do not have to play by the rules (Godkin & Allcorn, 2009), and silencing diverging minority opinions (Barton & Sutcliffe, 2009). Even beyond that, recognition of expertise can lead to a sense of entitlement or narcissism, that leads to envy when their goals are threatened, exceptional belief in own successes, driving out non-conformers, managing by intimidation, and suppressing accurate testing. The localisation of specialised knowledge often makes it hard for others to check on what the experts do. Especially in places of high ambiguity, prima donna's can assign credit to themselves, while diffusing blame away (Dekker, 2014a). This is counterproductive to other principles from HRO and Resilience Engineering such as being sensitive to operations and understanding the larger picture, aligning groups to work together across multiple levels, taking minority opinions seriously, and to not take past successes as a reason for confidence (Hollnagel et al., 2006; Weick & Sutcliffe, 2007).

5.6.7 Commitment to resilience

In HRO studies, resilience is about responding to disturbances and returning to normal operations. This is a narrower view than how resilience is used in Resilience Engineering, where it also includes prediction and prevention of future events. Nevertheless, the ability to recover from disturbances is an important element in any definition of resilience, and both HRO and Resilience Engineering value preparing for unexpected events.

As Resilience Engineering often aims to understand work, research often focuses on how emergency teams work comes together. This is then generally used in two ways, either to provide teams with more resources and better tools for the challenges they face, or to identify the elements that make up 'emergency response' expertise, so this can be used in training (Furniss, Back, Blandford, Hildebrandt, & Broberg, 2011). We will focus on the later, the elements identified from successful expert teams.

One of the most returning issues on emergency response is the need to plan, establish goals, and distribute roles (Gomes et al., 2014; Righi, Huber, Gomes, & de Carvalho, 2016). This is not to rigidly stick to those goals, as resilience is about creating the capacity to adapt. However, in emergency response teams these actions help to establish common ground to aid in coordination and to allow people to be proactive (Bergström, Dahlström, Dekker, & Petersen, 2011). In emergencies, it has been found that making modular plans works better, with sub-plans that are structured and integrated into larger plans. This adds to the ability to revise and re-organise quickly (Gomes et al., 2014), which is the next point.

In crisis operations it is expected that plans change and replanning and restructuring will be required. Modular plans make it easier, but it leaves open the question who is updated or when is a change significant enough to update everyone (Bergström et al., 2011)? From studying teams at work, it became clear that formal centralised (re)briefings spread information quickly and helped people know what others know. Wrong information that has been gathered also has the potential to spread quickly this way, but, during formal briefings created moments where wrong or implausible information was spotted (Rankin et al., 2013). Central updates, however, also take up time and impede work.

Adapted team structures often arise from what is necessary or possible, but might not be optimal in the long term. Teams identified that both bottom-up and top-down processes can prompt searches for more optimal structure after one pattern has stabilised (Lundberg & Rankin, 2014).

Re-structuring teams often leads people to work on tasks outside their own area of expertise (Bergström et al., 2011). While it will often be impossible to cross-train everyone in all skills, it can be possible to assist people in working beyond their expertise. This can include creating the expectations that role-improvisation will happen, teaching people how to support non-experts in their speciality, finding time during operation to prepare others working outside their expertise, as well as the challenges of communicating about a topic where one does not have expertise (Lundberg & Rankin, 2014).

At the team and management level, the expertise for resilience is largely in how teams self-organise and interact. Researchers studying emergency response teams has certainly been done before the emergence of Resilience Engineering and elements such as planning and maintaining a shared view have been recognised before and have been covered in CRM training (Flin, 1995).

Resilience Engineering has a slight shift of focus. In Resilience Engineering inspired research, the expertise is less seen as requirements people should meet, like skills of individuals, or things that should not happen, like avoiding 'errors'. With the view of Resilience Engineering, there are no more clear standard to meet that ensure safety, instead it is more phrased as things that help muddling through (J. M. J. Flach, 2012),.

5.6.8 Preoccupation with failure

HRO actively search for things that could go wrong (Weick & Sutcliffe, 2007). One way that organisations do this is by assessing risk. Risk assessment ranges from estimating the rate of occurrence of historically frequent events, to judging the likelihood of rare or even unprecedented high consequence events. For low frequency, high consequence events "expert judgement" is used to assess and prioritise scenarios. However, there is limited evidence that design or operational expertise translates into superior predictive ability (Rae & Alexander, 2017).

To the extent that some people make more accurate predictions than others, this appears to be expertise in translating qualitative impressions into quantitative estimates more generally, rather than a function of subject-matter knowledge or domain experience (Mellers et al., 2015).

However, the value of experts in these risk assessments is not necessarily the accuracy of their numbers, but also how they produce these numbers. When multiple experts decompose their estimates and discuss together what they have done, they often reach new conclusions (Rae & Alexander, 2017).

For the purpose of enhancing resilience, there is ongoing debate how best to identify and make use of subject matter expertise. One school of thought suggests that superior performance on calibration problems should be the sole determinant of expertise (Cooke & Goossens, 2004). The alternative is to consider quantification as a means rather than an end, providing a focus for discussion between stakeholders with domain expertise.

5.6.9 Concerns about expertise

As with individuals, people have raised the concerns that too much expertise of teams, as a unified agent, can lead to decreased attentiveness and sensitivity to operations. Weick and Roberts (1993) described this with the idea of heedful interrelating. Heedful interrelating is interacting with others in a way that others, the larger task, and the organisation, are taken into account, going beyond the local concerns. Repetition with the same people can lead to heedless habits or routine response, where each interaction is just a replication of what has been done before, instead of being modified by it. Weick and Roberts (1993) see the rotation of crew members and frequent introduction of new crew members on flight-deck carriers as a way crew members are forced to interact heedfully, instead of relying on routines. As routines build through experience, they represent a form of team expertise.

While the idea of heedful interrelation often comes up, there are few empirical studies on expertise and heedful interrelating, and whether it leads to better performance. Gorman and Cooke (2010) found in experiments where teams with three members controlling Uninhabited Air Vehicles that the teams that were kept intact were more corrective in their interaction patterns, while teams that where the team members were mixed around explored more solutions to encountered problems. However, they did not test whether this was mediated by heedful interrelating. Grote, Kolbe, and Künzle (2010) analysed the behaviour of cockpit crews in a simulator. They did not find that high performing cockpit crews used more heedful interrelating, but did spot qualitative differences in the heedful interrelating of high and low performing crews. They also found that heedful interrelating was more common in high workload tasks, and less common in more standardised tasks. Implicit coordination was higher in both high workload and high standardisation condition. This confirms the idea that heedful interrelating is different from routine team action. Schraagen (2011) analysed the interactions of one medical team during multiple paediatric cardiac surgeries and found that during non-routine events, the team mostly relied on explicit coordination, not heedful interrelation. However, in-depth qualitative analysis of the interactions suggested that heedful interrelations did form critical moments. The study concluded that heedful interrelating is important, but most non-routine events are dealt with through routine procedures.

More research needs to be done in this area. That rotating team members made teams more flexible, could have other explanations than changes in heedfulness. For example, it could a product of the diversity in known options within one team instead of heedfulness. In addition, routines are also used to deal with non-routine situations, suggesting teams always rely on a shared background and never only do heedful interrelating. This is supported by the idea that common ground is necessary for joint activity (G. A. Klein, Feltovich, et al., 2005). The expertise of a team on a whole appears to allow teams to respond faster with less costs of coordination, while diversity of expertise between team members appears to assist finding new options.

5.7 Systems level

Woods and Branlat (2010) identified three basic patterns through which adaptive systems, including organisations fail. These patterns are de-compensation, working at cross-purposes, and getting stuck in out-dated behaviours. We will discuss these three patterns, and how to manage expertise can help organisations move away from these patterns.

5.7.1 De-compensation

A system de-compensates when challenges or disturbances are stronger or faster than the system can adapt to. The system has to deal with too much in too little time. This can be a matter of ability, but also of resources it has available.

From a reactive perspective, expertise can increase the ability to cope with disturbances. The pitfalls people have associated with expertise, including fixation and routinisation, can even be the processes that provide this extra capacity, as these processes reduce the amount of information taken into account. However, expertise is only useful if the expertise and resources are available at the right place at the right time. It is not only about having the expertise, but also being able to access it. For accessibility it is beneficial to have many connections within an organisation (Carmeli, Friedman, & Tishler, 2013).

More connections within an organisation give local agents shorter paths to expertise and people that hold the resources. This is also seen on an individual level, where the most resilient individuals seek out relationships with people who hold the key resources (Lengnick-Hall & Beck, 2008). However, maintaining connections can be costly in a system, especially if they are rarely used, which is why networks are rarely fully connected.

From a proactive perspective, expertise can be used in the monitoring and prediction of larger disturbances. The schema and patterns of situations that experts have help experts to quickly spot things stand out from the norm (Feltovich et al., 1997; G. A. Klein, Pliske, et al., 2005). Expertise is also helpful for predicting how an anomaly influences a system, and considering what its possible causes are of a failure, and with that determine whether an anomaly is relevant (Watts, Woods, & Patterson, 1996). As discussed, this deeper understanding of how a system links together makes expertise useful for risk assessments, even when it does not increase accuracy.

5.7.2 Working at cross-purposes

Working at cross-purposes happens when subgroups are adaptive towards their own goals, but are maladaptive towards higher-level goals. This can happen when different groups draw from a common resource pool, such as communities of fishers (Ostrom, 1999), or when the work of one group affects the work of another group. Vaughan (1996) described how at NASA and Snook (2000) at military operations in Iraq how small changes to practices were locally adaptive, but caused misalignment on a larger level, leading to eroding safety margins and unintended outcomes.

Local adaption is overall considered useful, necessary, and even inevitable, in sociotechnical system. This shows in the advice to give decision freedom to frontline workers and to defer to expertise. The challenge on a systems level is to prevent local adaptation from becoming maladaptive at a larger level. There are three types of solutions to this problem:

- *Regulations:* To prevent lower groups from acting maladaptive to larger groups, it is possible to constraint their actions with rules or regulations. This will limit the resilience at a lower level, but preserve the system at a higher level.
- *Central controller:* One way to prevent maladaptation is to add a controller for the higher-level goals, overseeing the lower levels. An example of this in aviation, is air traffic control, who manage limited airport runway and air space and exert their influence on arriving and departing airplanes. For maritime ports, a similar centralised function has been suggested (Westrenen, 2014).
- *Adapting to adaptations:* The third option is to make the groups adapt to the adaptations of others. This relies on making the larger and shared goals and safety margins visible, as well as how other groups are behaving. Ostrom (1999) has found groups tend to cooperate and form their own norms and adjust for each other's behaviour, as long as there are enough face-to-face interactions within that community. Patterns that emerge from local adaptations can be just as efficient when people can see what is available and know how larger goals can be achieved (Hutchins, 1991; Ostrom, 1999).

All of these methods rely on a kind of expertise to manage the system. For the first experience is translated into rules. The second, a central controller, will need a model on how the agents it is trying to guide work (Conant & Ashby, 1970), which is usually built on experience and can continue to update with experience. In the third, the expertise is not centralised, but distributed over the agents in the system. The expertise takes a different shape with each approach.

5.7.3 Getting stuck in outdated behaviours

In this pattern, a system keeps doing what it has done in the past, even though the situation has changed and requires a different response now (Woods & Branlat, 2011). Cognitive inflexibility (Spiro et al., 1988) and routine actions (Weick & Roberts, 1993) represent this pattern on smaller scales. This pattern reflects a failure to learn, while paradoxically also being linked to the product of learning. An example of this pattern is organisations that stick to a narrow interpretation in their analysis of accidents, failing to produce new learning (Cook, Woods, & Miller, 1998).

To counter this effect on an organisational level, learning needs to be encouraged. Often this is linked to the idea of instilling a learning culture. An important element to learning is that learning opportunities are recognised and collected. Accidents and incidents, for example, provide learning opportunities, however, the learning can be lost if people are afraid of speaking up about the accident and they cannot be investigated further. This is why the learning in an organisation greatly suffers, if there is not a just culture that protects against blaming and shaming (Dekker, 2016). While this alone is not sufficient for organisation to learn from accidents, it is a first step to spread learning in an organisation (Huber, Wijgerden, Witt, & Dekker, 2009).

To prevent an organisation from getting stuck in dated behaviours, it is useful if the groups in an organisation are interconnected. Interconnectedness can help recognise change conditions and that safety margins are eroding. For example, an outsider who does not have a preconceived notion of the process or has the resources to take a larger view of the system, might spot that a stabilised task distribution is less than optimal in the current situation (Lundberg & Rankin, 2014) or how work in one area influences on other parts of the system (Dekker, 2011; Patterson et al., 2007).

For the challenge of finding new ways to act, diversity of agents can outperform the quality of single problem solvers. An individual with similar expertise will tend to converge to similar solutions. Here is where fresh perspectives add more to expertise of the group as a whole, because they are different (Hong & Page, 2004). However, diversity can increase coordination cost. People from diverse background can have a harder time working together as it will be harder to establish common ground (G. A. Klein, Feltovich, et al., 2005). This can slow a system down and put it at a risk of de-compensation.

5.8 Conclusion

When we look at expertise and resilience, we see a different relationship across the different levels. At the individual level of frontline workers, resilience and expertise are almost one and the same. More expertise is almost always more resilient. At the level of management and team expertise is still beneficial, but the type of expertise that is useful, how to make experts collaborate, and political implications of recognising experts begin to matter. At a systems level, it becomes harder to point to expertise, as it can take many shapes. The question becomes not whether to expertise is required, but what shape it is and how can it be accessed.

If we look at the larger picture of research on expertise and resilience, there is limited research going across scales (Bergström & Dekker, 2014), and despite that Resilience Engineering aims to understand work in general, research has mostly focussed on dealing with negative disturbances. How organisations deal with more positive disturbances is hardly addressed.

There is some work that mentions cross-scale work, such as the working at cross-purposes pattern, but there are few empirical studies on patterns across scales. For example, do diverse groups of individuals outperform experts in risk assessment? In assessment of low-frequency high-consequence events, the value of experts seems mostly in the links they make between different elements of the system and not in their accuracy of the chance. If accuracy is not the goal, diverse groups might be better at seeing connections then experts are. Both comparisons on the process as to how the risk assessment is made, as well as the outcome accuracy, can provide answers.

Resilience Engineering has largely focussed on the reactive part, rather than the foresight and proactive goal that was envisioned for resilience engineering. Resilience Engineering also encompasses creating and using foresights, leveraging extra available resources, and spreading insights through an organisation. There has been some mention of investigating how systems can flourish (Furniss et al., 2016), but the topic has not been explored yet. To further explore these topics, Resilience Engineering can build on research such as divergent thinking (McCrae, 1987) and insight (G. A. Klein & Jarosz, 2011), which both seem to require expertise, but also acknowledge that more is not always better (Runco, Dow, & Smith, 2006). As this would be explore a new area, explorative research is most likely to help us understand the challenge of ‘positive’ disturbances.

5.9 Outcome for the research aims

In terms of how to organise blast crews, this review led to conclusions on the level of individuals and systems.

For individual operators, autonomy is desirable. Having decision authority facilitates operators to develop and use expertise in the form of a deep problem understanding and multiple solution structures. Regarding team level routines, because of this it becomes desirable that there are few points of interaction, to give individuals more room to vary. In addition, I found confirmation in other studies that also discovered that repeated experience with a strategy helps with spotting deviations, which can help operators recognise things are not going as they expected. While domain expertise is beneficial to successful adaptation, there was little indication that there is a general expertise in dealing with unknown situations.

On the level of systems, there are arguments for what problem organisation by routines can be beneficial and for which not. Organisation by routines can help against decompensation, as it increases efficiency. By needing to spend less time on planning and explicit coordination, a team has more time to deal with what encounter. Of course, this is within the limitation of a specific routine, but the same pattern of activity performed by a crew who is new to it will require more effort, than for a crew that routinely uses that pattern.

Routines can also serve as a barrier against working cross-purposes. In blast crews, the individuals already have largely the same goals. However, having a shared solution structure allows teams to have the smaller steps combine efficiently together and prevents cross-purpose actions at the smaller levels. Between blast crews and other groups, routines could prevent working at cross-purposes by capturing a truce between competing goals (Becker et al., 2005), but this is not something I have observed in my fieldwork. Routines fulfil this function by providing implicit rules. However, these rules are not static and adapt with the development of expertise in teams.

In relation to getting stuck in out-dated patterns, organisation by routine actually enables this pattern. Within the freedom of the solution structure, individuals can vary, but the shared solution structure forms a pattern that might be too slow to change. There are conditions that can make a routine more fluid (or robust), which can decrease the chance of getting stuck in an out-dated pattern. Three possible intervention to make routines more fluid are 1) involving people with diverse experiences, 2) mixing up team membership, and 3) a just culture (Dekker, 2016) that allows people to talk about errors or things that might be considered deviant.

Chapter 6: **General discussion**

The discussion is split into the achievements of this thesis of its research aim, its theoretical and methodological contributions, and the practical applications of its research findings.

6.1 Achievement of research aim

The general aim of this thesis was to investigate how teams organise for safety, efficiency, and other goals, with a specific focus on blast crews in mining. This aim with the practical application in mind of improving the organisation of team and the organisation for teams in regards to safety. Safety here was approached as the ability to deal with varying circumstances, as is the focus in Resilience Engineering. This research aim did not come about because blast crews are considered unsafe or have high accident rates. Instead, it arose from a search for alternative ways of improving safety, as safety levels had reached a plateau and stopped improving. Improvement was not expected to come from new technologies or tighter control, but from a better understanding of the complexity of work, social construction, and the way work is organised.

Table 11 Research contribution and questions per chapter

Chapter	Contribution to research aim	Research Question
2	Review on prescribed ways of organising teams	How has CRM been adopted across industries?
3	Methodology	How to investigate everyday work for safety purposes?
4	Empirical study of how teams organise	How do teams organise routine work?
5	Extension of results through review	How to manage expertise to enhance resilience?

Chapter 2 was an investigation into the state-of-the-art way of organising teams for safety in other domains. As CRM is often described as highly successful, this made me believe that it likely contained useful ideas on how teams organise or how to organise team. However, my investigation did not return such results.

The review of the CRM literature showed me that programs under the CRM are too diverse and there are too few meaningful evaluative studies conclude from what is a good way to organise teams for safety. There was very little evidence that suggested any CRM program has been effective in the reviewed industries. With so few successes and so much diversity in what programs aim to do, I could not distill lessons on what kind of organising would work for blast crews. The investigation did show that there is a lack of evidence to say that teaching teams to organise is effective for improving safety. In addition, the review pointed to gaps in the literature, or at least a lack of consensus, on the way teams organise, how team organisation comes about, and what influences how teams organise. These gaps are partly addressed in the subsequent chapters.

Chapter 3 forms a methodology chapter, where I asked how to investigate the everyday work for safety. This larger reflection on how to investigate everyday work for safety guided the design of the empirical research. The review found a general tension between focussing on safety, the prevention of specific rare events, and the everyday, the unspecific and ubiquitous. For normative and positivist research, which usies units of analysis based on accident model concepts, it is easier to establish the relevance to safety, but might lead to a view that fits poorly with what everyday work is like. In descriptive and constructivist approaches, and with units of analysis not based on accident models, it becomes easier to learn something new about the complexity of everyday work (even though it is less certain beforehand how a topic relates to safety).

Based on the CRM review, and the small number of studies on blast crews, there was not a lot known to be relevant to blast crews and safety from academic sources. This lack of guidance led to a descriptive approach, with constructivist elements, which allowed me to explore the unknown elements in how blast crews organise work. In the formulation of my original research aim, I had already built on the HRO literature and the relation they established between organising and safety.

Chapter 4 presents the major empirical work of this thesis. With a problem-orientated ethnography on blast crews and linesmen, I investigated how teams organise their work safely and efficiently through routines. Building on theories from CSE and organisational routines, I explain how individual decisions give rise to general repeated patterns of work at a team level. I found that operators have a stable problem understanding of their task, which influences their work and the mutual coordination for it. There are also recognisable patterns in how a team approaches a task. In contrast to how the organisational routines literature describes it, I found that teams are not trying to repeat patterns of actions they have done before. Instead, individual operators have a repetitive way in which they divide and structure a larger task into smaller tasks, which I have called the solution structure. When this solution structure is shared across team members and specified to the task and conditions, this leads to a smooth and coordinated process with which a team completes their work. The problem understanding and solution structure allow crew members to give meaning to the environment, in terms of recognisable signs. This repetition in structure makes people more sensitive to the operations, since both intentional and unintentional deviations stand out.

Our results explain how blast crews organise routine work. This explanation can be used as an example for how to organise teams for safety, particularly because blast crews are generally considered to be quite safe. However, considering that this is how blast crews self-organise in ways that arose without intervention or top-down guidance, it is likely of lesser value as a teaching tool for others. I have, however, let the discoveries from the empirical work inspire my search for some of the key ingredients that enable this type organising. One such ingredient, expertise, was addressed in Chapter 5.

Chapter 5 provides a review on how to manage expertise across different levels to enhance resilience, the type of safety valued in this thesis. At the individual level, more expertise generally means more resilience. Because of this, individuals should be given decision autonomy so they can deploy their expertise *and* develop their expertise. There are, however, challenges in how to combine multiple experts. and multiple goals and concerns. At the systems level, expertise can help against multiple ways of system breakdown. The challenge here is not necessarily having enough expertise, but having the right type and mix of expertise and having it accessible when it is needed. Diversity in expertise across agents can be as useful as in-depth expertise.

In relation to how blast crews organise routine work, expertise can be leveraged in multiple ways for resilience. In a routine, a solution structure with fewer points of interactions gives individuals more freedom to vary, which gives them more opportunity to develop and use expertise. On the team level, keeping teams together can build robust routines, while rotating team members can make routines more fluid by increasing the diversity in expertise within a team and prompting searches for new routines. At a system level, organising by routines can help against de-compensation, as it increases efficiency. Routines can also serve as a barrier against working at cross-purposes, as they provide implicit rules to align behaviour towards a larger picture, which limits the possible negative effects of local adaptions. In relation to adaptive systems getting stuck in out-dated patterns, routines describe a method in which this can actually happen. At an organisational level, this can be addressed by having people that affect, and are affected by, a team reflecting on the current process.

6.2 Scientific contribution

6.2.1 Theoretical contribution

This thesis makes theoretical contributions about CRM, dealing with complexity, and organisational routines.

CRM

The review of CRM adoption across industries has shown that there is little evidence that supports the efficacy of CRM for non-pilots. On top of that, there is not one uniform conceptualisation of CRM. This variation in conceptualisation means that just doing more evaluation is not going to answer the question as to CRM is doing what people claim it does. To assess whether CRM does what people say it does, the conceptualisation of CRM first needs to be made clearer. However, considering the long history of CRM and existing diversity, I consider it infeasible to move forward with one standardised conceptualisation. This is why I argue that in each study, researchers need to make explicit the process through which they expected CRM to lead to changes.

This theoretical finding could have consequences for how CRM training is conducted. Data suggest that the assumed model of individual learning in terms of knowledge, skills and attitude, leading to changes in behaviour and, in turn, improving safety and efficiency, does not fit. The lack of confirmation of that model could mean that where CRM has been found to have an effect on safety and efficiency, that this happened via another process. More research is needed to verify this, but if CRM actually works via different processes, this could be leveraged in the design of CRM training.

Dealing with complexity

As done in HRO, Resilience Engineering, and CSE work, by studying people at work, I discovered a pattern of how people deal with complexity. However, in contrast to the typical conclusions of HRO and Resilience Engineering literature, I conclude that teams use repetitions and stability in certain areas to deal with complexity in a coordinated way. The solution structure creates a modularised division of the problem, acting like levels of abstraction. Instead of countering variability, the discovered process shows how people try to reduce the variability. In agreement with HRO and RE, both within this structure, and some cases because of this structure, people can be adaptive, which is not only required for successful outcomes but is also more efficient.

Our results challenge the distinction between heedful interrelating and routine interaction as drawn up by Weick and Roberts (1993) in the HRO literature. I found that even for routine actions, individuals relate the current situation to a bigger picture. Routine actions cannot be explained as replicating responses to known signs. People can differ in how far they look ahead or how big of a picture is considered, but if anything, expertise and team member stability increase the degree to which the bigger picture is taken into account. This is opposed to the idea that team member rotation increases heedfulness (Weick & Roberts, 1993). My results show that routines enable people to be sensitive to operation. However, in line with Weick and Roberts' (1993) statements on heedful interrelating, my conclusions suggest that it is plausible that routines, or solution structures, are more likely to change with the rotation of team members. Rotation of team members leads to divergent problem understanding and solution structures within a team, which can prompt a search for a fitting routine.

In the CSE literature, Rasmussen (1983) recognised people as teleological by nature. In his skill, rules, and knowledge framework, he described skill and rule decisions — the decisions that rely on repeated experience — as goal-orientated. My explanation adds how team processes can shape the goals of such individual decisions.

The Einstellung effect (Luchins, 1942) and problem restructuring (Duncker & Lees, 1945) suggest that the idea of a solution structure and problem understanding also apply when people are working alone. However, as team processes influence the solution structure in the division, structure, and acceptable in-between-states, this leaves the question of how this develops differently for individual and team tasks. It is imaginable that teams can both resist and facilitate change in solutions structures. Teams could resist change as more people will be affected and need to adjust to any changes, thus increasing the cost of coordination. Teams could also facilitate changes in solution structure, as teams bring different perspectives together and with more people around it increases the variation individuals encounter, which can prompt change. Further research is needed to explore this question.

Organisational routines

Similar to the organisation routines literature, I found stability in how groups approach a task, which can remove the need for planning and explicit coordination. However, on multiple points my conclusions contradict or extend the organisational routines literature. My observations contradict the organisational routines literature in that I did not find repeating ‘patterns of actions’. Instead, I found that people repeat the way they divide and structure a problem into smaller tasks, the solution structure. This solution structure does not specify actions in detail, but it does provide the goals to move towards. The actions then come about from the goals and the possibilities the situation provides. This change from actions to solution structure challenges the definition of organisational routines, when used to describe routines from an agent perspective. From this also follows that routines cannot be decomposed to infinite in detail.

In addition, I show that practitioners do not only rely on a solution structure or ostensive image, but also rely on a problem understanding to read their environment and produce actions. The combination of a problem understanding and a solution structure of the individuals in a team explains how 1) different actions in a routine link together, 2) with experience, people can become more adaptive, rather than more rigid, as the idea of imprinting actions suggests, and 3) routines coalesce as a team works together for a longer period.

6.2.2 Methodological contribution

I provide two methodological contributions for the study of everyday work for safety and for the study of changing routines.

Investigations of everyday work for safety

I provided a reflection that can aid the design of everyday work investigation for the purpose of safety. One thing I identified was the challenge of linking the everyday to safety. Without an accident as a starting point, there needs to be a case made as to why something found in ubiquitous everyday work is relevant to safety, which is generally linked to special cases. One way to do this is by looking for elements found in accident models. Another way, as proposed by HRO and Resilience Engineering, is to look at the flip side of accidents, which is success.

Changing routines

I established the concept of solution structure to explain what does and does not repeat between multiple instances of the same routine. With this, it becomes possible to recognise what is 'daily' variation and what change comes from changing routines. The existence of this difference exists has been recognised in the organisational routines literature (Becker et al., 2005), but to my knowledge it has never been accounted for in the methods of any previous study. By capturing the solution structure at multiple instances, it becomes possible to separate the 'noise' of variation within a routine, from variation of interest that comes from changing routines. This can also be applied to research on (procedural) drift (Dekker, 2011; Snook, 2000) and normalisation of deviance (Vaughan, 1996).

Currently, there are no standardised methods to uncover solution structures, but the methods used in chapter 4, as well as other cognitive task analysis techniques (Crandall et al., 2006), can provide a starting point for future research.

6.3 Practical implications

As done in CRM, it is possible to use the explanation of how routine work is organised in training, so that others can apply it. However, since this type of organising developed without intervention, and the organisational routines literature suggests it is common, I think there is little value in this approach. To assist the organisation of routine work, I think it is of more value to create conditions beneficial to the organisation of routine work. Two ways that organising routine work can be assisted are through team composition and procedures.

6.3.1 Team composition

As was originally raised with the idea of heedful interrelating, team composition and team member rotation can influence existing routines. My results challenge the idea that routines are not heedful, but acknowledge that having an effective routine can prevent a search for a better one. Rotating team members can prompt team members to search for new and better routines. I describe this with the idea of robustness and fluidness of a routine.

A team that deals with changing conditions and has stable team membership will develop robust routines. This means the team members have a shared problem understanding and solution structure, which makes them interpret the signs similarly and have similar responses to disturbances, without explicit coordination. This reduces the cost of coordination and allows them to respond quickly to disturbances. Robustness is generally desirable; however, it makes it unlikely a routine will change beyond its current solution structure. The robustness can lead a team to stay with a routine, even when better solutions are possible.

When team composition changes, routines become more fluid. With changes to team composition, there is no one assumed solution structure in the team anymore, and there can even be different problem understandings across team members. This forces teams to think about how they divide and structure the task into smaller task and reflect on the problem they are trying to solve. Initially, this will increase the cost of coordination and planning, and make individual team members more uncertain of the meaning of signs in the environment and how to respond to them. However, in most cases teams will move towards a routine that fits the current conditions.

Generally, robustness is preferable, which aligns with the finding that team members that are familiar with each other generally perform better than newly created teams (Foushee, Lauber, Baetge, & Acomb, 1986). However, this robustness can also be a detrimental factor, in that it limits the solutions considered. This aligns with the idea that changing team membership can make teams more adaptive (Gorman et al., 2010). From this follows that team stability is preferred, with occasional rotations. Creating robust routines in general, with moments of fluidity to adjust to possibly changed conditions. The amount of variation encountered, as well as the predictability of the type of variation, affect whether more robustness or fluidity is desirable.

In addition, when creating a new team, combining people with similar solution structures will create a robust team more quickly, while combining people with diverging solution structures will likely to lead to a wider search for solutions. This suggests that when a team needs to perform well quickly, people with similar problem understandings and solution structures are preferred. If the new team needs to work in unfamiliar conditions, it is preferable to combine people with diverging types of experience in the team.

6.3.2 Procedures

A shared problem understanding and solution structure aid teams in dealing with complexity in multiple ways. It aids coordination in that it allows people to interpret the current situation, helps them in what to expect, and informs people on what they need to do, while still giving them freedom to adapt. On top of that, it can prevent different agents from working at cross-purposes to each other. While I do not think it is useful to teach teams to do this, the concept can be incorporated and mimicked when writing procedures.

Generally, it is beneficial to minimise procedures and to give individuals more autonomy, which allows them to develop and use more expertise. For this you want to reduce the amount of procedures as much as possible. However, when people lack the expertise to make their own decisions, more guidance on how to achieve a task can be beneficial and procedures can avert people working at cross-purposes.

When procedures are used to train novices, it can be helpful to communicate how to achieve goals, and divide a problem into smaller steps, as they might not know how to achieve larger steps at once. However, these procedures should not be monitored on compliance more than necessary, as experimentation is desirable to build expertise. The goal is to have novices develop a problem understanding that allows them to understand what is and is not possible. This understanding needs to be developed in relation to their local environment and the big picture of the organisation.

To prevent working at cross-purposes, while minimising any decrease of autonomy, the concept of solution structure can provide guidance. The solution structure can inform the level of specification needed in procedures. Procedures used by experienced groups should focus on points of interaction with other groups but can leave freedom on how to achieve those states. Groups will normally take care of this naturally, but when the different agents do not meet, cannot give feedback to each other, are not familiar with each other's work, or have unequal power relations, it can be useful to formalise the points of interaction. By keeping the goals standardised, but leaving the process undefined, individuals can adapt to the circumstances, while the different pieces still combine to successfully complete a task.

References

Annett, J. (2000). Theoretical and pragmatic influences on task analysis methods. In J. M. Schraagen, S. Chipman, & V. Shalin (Eds.), *Cognitive task analysis* (pp. 25–37). Mahwah, NJ: Lawrence Erlbaum Associates.

Ashby, W. R. (1957). *An Introduction to Cybernetics*. London, UK: Chapman & Hall ltd.

Avermaete, J. A. G. van, & Kruijsen, E. A. . (1998). *NOTECHS: The evaluation of non-technicals skills of multi-pilot aircrew in relation to the JAR-FCL requirements*. *Final Report NLR-CR- 98443.* Amsterdam, The Netherlands.

Bailey, T. (2014). BP uses high-fidelity simulators to develop non-technical skills. *Offshore Engineer*, *39*(4), 32–34.

Bainbridget, L. (1983). Ironies of Automation*. *Automatica*, *19*(6), 775–779.

Barbarino, M., & Isaac, A. (2000). Development of Team Resource Management in European air traffic control. In A. R. Lowe & B. J. Hayward (Eds.), *Proceedings of the Fourth Australian Aviation Psychology Symposium* (pp. 269–278). Aldershot, UK, UK: Ashgate.

Barnett, M. L., Gatfield, D., & Pekcan, C. (2003). A Research Agenda in Maritime Crew Resource Management. In *Proceedings of the International Conference on Team Resource Management in the 21st Century.* Embry-Riddle Aeronautical University. Retrieved from http://ssudl.solent.ac.uk/433/1/Embry-20Riddle-202003.pdf

Barsan, E., & Hanzu-Pazara, R. (2010). Human response in navigation emergency situation. *International Journal of Vehicle Safety*, *5*(1), 75–85. http://doi.org/10.1504/IJVS.2010.035320

Barton, M. a., & Sutcliffe, K. M. (2009). Overcoming dysfunctional momentum: Organizational safety as a social achievement. *Human Relations*, *62*(9), 1327–1356. http://doi.org/10.1177/0018726709334491

Becker, M. C. (2004). Organizational routines: A review of the literature. *Industrial and Corporate Change*, *13*(4), 643–677. http://doi.org/10.1093/icc/dth026

Becker, M. C., Lazaric, N., Nelson, R. R., & Winter, S. G. (2005). Applying organizational routines in understanding organizational change. *Industrial and Corporate Change*, *14*(5), 775–791. http://doi.org/10.1093/icc/dth071

Bergström, J., Dahlström, N., Dekker, S. W. A., & Petersen, K. (2011). Training Organisational Resilience in Escalating Situations. In E. Hollnagel, J. Pariès, D. D. Woods, & J. Wreathall (Eds.), *Resilience Engineering in Practice: A Guidebook* (pp. 45–57). Farnham: Ashgate.

Bergström, J., & Dekker, S. W. A. (2014). Bridging the Macro and the Micro by Considering the Meso: Reflections on the fractal nature of resilience. *Ecology and Society*, *19*(4).

Bergström, J., van Winsen, R., & Henriqson, É. (2015). On the rationale of resilience in the domain of safety: A literature review. *Reliability Engineering & System Safety*, 1–11. http://doi.org/10.1016/j.ress.2015.03.008

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts-Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology*, *56*(2), 73–102. http://doi.org/10.1016/j.cogpsych.2007.02.001

Bourrier, M. (2011). The Legacy of the High Reliability Organization Project. *Journal of Contingencies and Crisis Management*, *19*(1), 9–13. http://doi.org/10.1111/j.1468-5973.2010.00628.x

Bradley, B. H., Postlethwaite, B. E., Klotz, A. C., Hamdani, M. R., & Brown, K. G. (2012). Reaping the benefits of task conflict in teams: The critical role of team psychological safety climate. *Journal of Applied Psychology*, *97*(1), 151–158. http://doi.org/10.1037/a0024200

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Brun, W., Eid, J., Johnsen, B. H., Laboerg, J. C., Ekornas, B., & Kobbeltvedt, T. (2005). Bridge Resource Management Training Enhancing Shared Mental Models and Task Performance. In H. Montgomery, R. Lipshitz, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 183–193). Mahwah, NJ: Lawrence Erlbaum Associates.

Buljac-Samardzic, M., Dekker-van Doorn, C. M., van Wijngaarden, J. D. H., & van Wijk, K. P. (2010). Interventions to improve team effectiveness: a systematic review. *Health Policy (Amsterdam, Netherlands)*, *94*(3), 183–95. http://doi.org/10.1016/j.healthpol.2009.09.015

Byrdorf, P. (1998). Human factors and crew resource management: An example of successfully applying the experience from CRM programmes in the aviation world to the Maritime World. In *23rd Conference of the European Association for Aviation Psychology, Vienna, Austria.*

Carayon, P., Alvarado, C. J., & Schoofs Hundt, A. (2007). Work design and patient safety. *Theoretical Issues in Ergonomics Science*, *8*(5), 395–428. http://doi.org/10.1080/14639220701193157

Carim, G. C., Saurin, T. A., Havinga, J., Rae, A., Dekker, S. W. A., & Henriqson, É. (2016). Using a procedure doesn't mean following it: A cognitive systems approach to how a cockpit manages emergencies. *Safety Science*, *89*, 147–157. http://doi.org/10.1016/j.ssci.2016.06.008

Carmeli, A., Friedman, Y., & Tishler, A. (2013). Cultivating a resilient top management team : The importance of relational connections and strategic decision comprehensiveness. *Safety Science*, *51*(1), 148–159. http://doi.org/10.1016/j.ssci.2012.06.002

Carvalho, P. V. R., Santos, I. L. dos, & Vidal, M. C. R. (2006). Safety implications of cultural and cognitive issues in nuclear power plant operation. *Applied Ergonomics*, *37*, 211–223. http://doi.org/10.1016/j.apergo.2005.03.004

Cilliers, P. (2005). Complexity, Deconstruction and Relativism. *Theory, Culture & Society*, *22*(5), 255–267. http://doi.org/10.1177/0263276405058052

Cohen, M. (2007). Reading Dewey: Reflections on the Study of Routine. *Organization Studies*, *28*(5), 773–786. http://doi.org/10.1177/0170840606077620

Conant, R. C., & Ashby, W. R. (1970). Every good regulators of a systemmust be a modle of that system. *Int. J. Systems Sci.*, *1*(2), 89–97.

Cook, R. I., & McDonald, J. S. (1988). Cognitive Tunnel Vision in the operating room - analysis of cases using a frame model. *Anesthesiology*, *69*(3A), 497.

Cook, R. I., & Woods, D. D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 255–310). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, R. I., Woods, D. D., & Miller, C. (1998). *A Tale of Two Stories : Contrasting Views of Patient Safety*. Chicago, IL.

Cooke, R. M., & Goossens, L. H. J. (2004). Expert judgement elicitation for risk assessments of critical infrastructures. *Journal of Risk Research*, *7*(6), 643–656. http://doi.org/10.1080/1366987042000192237

Cooper, G. E., White, M. D., & Lauber, J. K. (Eds.). (1979). Resource management on the flight deck. In *Resource Management on the Flight Deck*. Moffett Field, California: NASA.

Cornelissen, M., McClure, R., Salmon, P. M., & Stanton, N. A. (2014). Validating the Strategies Analysis Diagram: Assessing the reliability and validity of a formative method. *Applied Ergonomics*, *45*(6), 1484–1494. http://doi.org/10.1016/j.apergo.2014.04.010

Cornelissen, M., Salmon, P. M., Jenkins, D. P., & Lenné, M. G. (2012). A structured approach to the strategies analysis phase of cognitive work analysis. *Theoretical Issues in Ergonomics Science*, *14*(6), 546–564. http://doi.org/10.1080/1463922X.2012.668973

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working Minds*. London: MIT Press.

Crichton, M. T. (2009). Improving team effectiveness using tactical decision games. *Safety Science*, *47*(3), 330–336. http://doi.org/10.1016/j.ssci.2008.07.036

Crichton, M. T., & Flin, R. H. (2004). Identifying and training non-technical skills of nuclear emergency response teams. *Annals of Nuclear Energy*, *31*(12), 1317–1330. http://doi.org/10.1016/j.anucene.2004.03.011

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, N.J.: Prentice-Hall.

Davis, J. (2001). Enhancing human peformance in the control room: sharing the experiences of British Energy. In *People in control: An international conference on Human interfaces in control rooms, cockpits and command centres* (p. 345).

de Keyser, V., & Piette, A. (1970). ANALYSE DE L'ACTIVITÉ DES OPÉRATEURS AU TABLEAU SYNOPTIQUE D'UNE CHAINE D'AGGLOMÉRATION. *Le Travail*

Humain, *33*(3), 341–352.

de Keyser, V., & Woods, D. D. (1990). Fixation Errors: Failures to revise situation assessment in dynamic and risky systems. In A. G. Colombo & A. S. de Bustamante (Eds.), *Systems reliability assessment* (pp. 231–251). http://doi.org/10.1007/978-94-009-0649-5

Debono, D. S., Greenfield, D., Travaglia, J. F., Long, J. C., Black, D., Johnson, J., & Braithwaite, J. (2013). Nurses' workarounds in acute healthcare settings: a scoping review. *BMC Health Services Research*, *13*(1). http://doi.org/10.1186/1472-6963-13-175

Deboo, C. K. N. (n.d.). *Maritime Resource Management - Based on the world renowed SAS-BRM Course*.

Dekker, S. W. A. (2003). Failure to adapt or adaptations that fail: Contrasting models on procedures and safety. *Applied Ergonomics*, *34*(3), 233–238. http://doi.org/10.1016/S0003-6870(03)00031-0

Dekker, S. W. A. (2005a). Past the edge of chaos. *Human Factors and Aerospace Safety*, *5*(3), 267–278.

Dekker, S. W. A. (2005b). *Ten Questions About Human Error: A new view of human factors and system safety*. London, UK: Lawrence Erlbaum Associates.

Dekker, S. W. A. (2011). *Drift into failure: From Hunting Broken Components to Understanding Complex Systems*. Farnham, England: Ashgate.

Dekker, S. W. A. (2014a). Deferring to expertise versus the prima donna syndrome: a manager's dilemma. *Cognition, Technology & Work*, *16*(4), 541–548. http://doi.org/10.1007/s10111-014-0284-0

Dekker, S. W. A. (2014b). *The field guide to understanding "human error."* Farnham, UK: Ashgate publishing Co.

Dekker, S. W. A. (2016). *Just Culture: Restoring Trust and Accountability in Your Organization* (3rd ed.). Boca Raton, FL, USA: CRC Press.

Dekker, S. W. A., Cilliers, P., & Hofmeyr, J.-H. (2011). The complexity of failure: Implications of complexity theory for safety investigations. *Safety Science*, *49*(6), 939–945. http://doi.org/10.1016/j.ssci.2011.01.008

Dekker, S. W. A., Nyce, J. M., & Hoffman, R. R. (2003). From contextual inquiry to designable futures: what do we need to get there? *IEEE Intelligent Systems*, *18*(2), 74–77. http://doi.org/10.1109/MIS.2003.1193660

Dekker, S. W. A., & Pruchnicki, S. (2013). Drifting into failure: theorising the dynamics of disaster incubation. *Theoretical Issues in Ergonomics Science*, *15*(6), 534–544. http://doi.org/10.1080/1463922X.2013.856495

Dekker, S. W. A., & Woods, D. D. (2010). The High Reliability Organization perspective. In E. Salas & D. E. Maurino (Eds.), *Human Factors in Aviation* (2nd ed., pp. 123–144). Burlington, MA: Academic Press.

Dijkhuizen, H. A., Butter, P. den, & Koning, B. R. (1996). Instruction, training and expertise advancement of pilots: The key elements in obtaining the highest possible quality of certified pilots in the Netherlands. In M. S. Chislett (Ed.), *Marine Simulation and Ship Manoeuvrability* (pp. 125–143). Rotterdam, NL: Balkema.

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. *Report No. ORC 80-2*. Berkley, CA.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological Monographs*, *58*(5).

Dunham, R. C., & Lützhöft, M. (2014). Interim Results of a longitudinal study into the perceptions of Bridge Team Management (BTM) in pilotage waters as experienced by Deck Officers studying at the Australian Maritime College. In *15th Annual gneral assembly International Association of maritime Universities* (pp. 118–124).

Edkins, G. D. (2002). A review of the benefits of aviation human factors training. *Human Factors and Aerospace Safety*, *31*(1), 247–273.

Fadde, P. J., & Klein, G. A. (2010). Deliberate performance: accelerating expertise in natural settings. *Performance Improvement*, *49*(9), 5–14. http://doi.org/10.1002/pfi

Farrington-Darby, T., & Wilson, J. R. (2006). The Nature of Expertise: A Review. *Applied Ergonomics*, *37*(1), 17–32. http://doi.org/10.1016/j.apergo.2005.09.001

Feldman, M. S. (2000). Organizational Routines as a Source of Continuous Change. *Organization Science*, *11*(6), 611–629.

Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing Routines

Organizational. *Administrative Science Quarterly*, *48*(1), 94–118. http://doi.org/10.1038/nn1246

Feldman, M. S., & Pentland, B. T. (2008). Issues in empirical field studies of organizational routines. In M. C. Becker (Ed.), *Handbook of Organizational Routines* (pp. 281–300). Cheltenham, UK: Edward Elgar Publishing Limited. http://doi.org/10.1093/icc/dth070

Feldman, M. S., Pentland, B. T., Adderio, L. D., & Lazaric, N. (2016). Beyond Routines as Things: Introduction to the Special Issue on Routine Dynamics. *Organization Science Publication*, *27*(3), 505–513.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1997). Issues of expert flexibility in contexts characterized by complexity and change. In *Expertise in context: Human and machine* (pp. 125–146). MIT Press.

Flach, J. M. J. (2012). Complexity: Learning to muddle through. *Cognition, Technology and Work*, *14*(3), 187–197. http://doi.org/10.1007/s10111-011-0201-8

Flach, J. M., & Voorhorst, F. A. (2016). *What Matters? Putting Common Sense to Work*. Ohio, USA: Wright State University Libraries. http://doi.org/10.1097/00001888-200102000-00027

Flin, R. H. (1995). Crew resource management for teams in the offshore oil industry. *Journal of European Industrial Training*, *19*(9), 23–27.

Flin, R. H., Martin, L., Goeters, K., Hörmann, H., Amalberti, R., & Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Human Factors and Aerospace Safety*, *3*(2), 95–117. Retrieved from http://www.abdn.ac.uk/iprc/uploads/files/NOTECHS HFAS proof copy.pdf

Flin, R. H., & O'Connor, P. (2001). Applying Crew Resource Management on Offshore Oil Platforms. In E. Salas, C. a. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: applications of resource management training* (pp. 217–234). Mahwah, NJ: Lawrence Erlbaum Associates.

Flin, R. H., O'Connor, P., & Crichton, M. T. (2008). *Safety at the sharp end*. Farnham, England: Ashgate Publishing limited.

Flin, R. H., O'Connor, P., Mearns, K., & Gordon, R. P. E. (1999). Crew Resource

Management for Offshore Production and Maintenance. In *Offshore Europe Oil and Gas Exhibition and Conference*. Aberdeen, Scotland: Society of Petroleum Engineers.

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 219–245. http://doi.org/10.1177/1077800405284363

Fonne, V., & Fredriksen, O. (1995). Resource management and crew training for HSV-navigators. In R. S. Jensen & L. A. Rakovan (Eds.), *Proceedings of the 8th International Symposium of Aviation Psychology* (pp. 585–590). Columbus, OH.

Foushee, H. C., Lauber, J. K., Baetge, M. M., & Acomb, D. B. (1986). *Crew factors in flight operations: III. The operational significance of exposure to short-haul air transport operations*. Moffett Field, California. Retrieved from http://hdl.handle.net/2060/19870003689

Furniss, D., Back, J., Blandford, A., Hildebrandt, M., & Broberg, H. (2011). A resilience markers framework for small teams. *Reliability Engineering and System Safety*, *96*, 2–10. http://doi.org/10.1016/j.ress.2010.06.025

Furniss, D., Curzon, P., Blandford, A., & Furniss, D. (2016). Using FRAM beyond safety: a case study to explore how sociotechnical systems can flourish or stall. *Theoretical Issues in Ergonomics Science*, *17*(5–6), 507–532. http://doi.org/10.1080/1463922X.2016.1155238

Gaddy, C. D., & Wachtel, J. (1992). Team skill training in nuclear power plant operations. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 379–396). Norwood, NJ: Ablex.

Garmezy, N. (1971). Vulnerability research and the issue of primary prevention. *Amer. J. Othopsychiat.*, *41*(1), 101–116.

Garmezy, N., & Streitman, S. (1974). Children at risk: The search for the antecedents of schizophrenia, part I. Conceptual models and research methods. *Schizophrenia Bulletin*, *1*(8), 125. http://doi.org/10.1093/schbul/1.8.14

Garnerin, P., Arès, M., Huchet, A., & Clergue, F. (2008). Verifying patient identity and site of surgery: improving compliance with protocol by audit and feedback. *Quality & Safety in Health Care*, *17*, 454–458. http://doi.org/10.1136/qshc.2007.022301

Girard, N. J. (2005). Dealing with perioperative prima donnas in your OR. *AORN*, *82*(2), 187–189.

Godkin, L., & Allcorn, S. (2009). Institutional narcissism, arrogant organization disorder and interruptions in organizational learning. *The Learning Organization*, *16*(1), 40–57. http://doi.org/10.1108/09696470910927669

Gomes, J. O., Borges, M. R. S., Huber, G. J., & Carvalho, P. V. R. (2014). Analysis of the resilience of team performance during a nuclear emergency response exercise. *Applied Ergonomics*, *45*(3), 780–788. http://doi.org/10.1016/j.apergo.2013.10.009

Gomes, J. O., Woods, D. D., Carvalho, P. V. R., Huber, G. J., & Borges, M. R. S. (2009). Resilience and brittleness in the offshore helicopter transportation system: The identification of constraints and sacrifice decisions in pilots' work. *Reliability Engineering and System Safety*, *94*(2), 311–319. http://doi.org/10.1016/j.ress.2008.03.026

Gorman, J. C., Amazeen, P. G., & Cooke, N. J. (2010). Team Coordination Dynamics. *Nonlinear Dynamics, Psychology, and Life Sciences*, *14*(3), 265–289.

Grinde, T. A. (1994). Emergency Resource Management training. In *SPE Health, Safety and Environment in Oil and Gas Exploration and Production Conference* (pp. 413–417). Jakarta, Indonesia: Society of Petroleum Engineers.

Grote, G., Kolbe, M., & Künzle, B. (2010). Adaptive coordination and heedfulness make better cockpit crews. *Ergonomics*, *53*(2), 211–228. http://doi.org/10.1080/00140130903248819

Grote, G., Weichbrodt, J. C., Günter, H., Zala-Mezö, E., & Künzle, B. (2009). Coordination in high-risk organizations: The need for flexible routines. *Cognition, Technology and Work*, *11*(1), 17–27. http://doi.org/10.1007/s10111-008-0119-y

Gunderson, L. H. (2010). Ecological and Human Community Resilience to Natural Disasters. *Ecology and Society*, *15*(2), 18. http://doi.org/18

Gunderson, L. H., & Holling, C. S. (Eds.). (2002). *Panarchy: Understanding transformations in human and natural systems*. washington: Island press.

Haavik, T. K., Antonsen, S., Rosness, R., & Hale, A. (2016). HRO and RE: A pragmatic perspective. *Safety Science*. http://doi.org/10.1016/j.ssci.2016.08.010

Haavik, T. K., Kongsvik, T., Johan, R., Olgard, J., Røyrvik, D., Bye, R. J., … Almklov, P. G. (2017). Johnny was here: From airmanship to airlineship. *Applied Ergonomics*, *59*, 191–202. http://doi.org/10.1016/j.apergo.2016.08.028

Habberley, J. S., Barnett, M. L., Gatfield, D., Musselwhite, C., & McNeil, G. (2001). *Simulator Training for Handling Escalating Emergencies*. Southhamption.

Hackman, J. R. (1993). Teams, Leaders, and Organizations: New Directions for Crew-oriented Flight Training. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit Resource Management* (pp. 47–69). San Diego, CA: Academic Press.

Hale, A. R., & Swuste, P. (1998). Safety rules: procedural freedom or action constraint? *Safety Science*, *29*, 163–177.

Haller, G., & Stoelwinder, J. (2013). Is "Crew Resource Management" an Alternative to Procedure-Based Strategies to Improve Patient Safety In Hospitals? In C. Bieder & M. Bourier (Eds.), *Trapping Safety into Rules: How Diserable or Avoidable is Proceduralization? (ePub version)*. Farnham, England: Ashgate publishing.

Harrington, D. K., & Kello, J. E. (1991). Systematic evaluation of nuclear operator team skills training. *Transactions of the American Nuclear Society*, *64*, 182–183.

Harrington, D. K., & Kello, J. E. (1992). Systematic evaluation of nuclear operator team skills training: a progress report. In *Conference Record for 1992 IEEE Fifth Conference on Human Factors and Power Plants* (pp. 370–373). Monterey, CA, USA: IEEE.

Hassall, M. E., & Sanderson, P. M. (2014). A formative approach to the strategies analysis phase of cognitive work analysis. *Theoretical Issues in Ergonomics Science*, *15*(3), 215–261. http://doi.org/10.1080/1463922X.2012.725781

Havinga, J., Dekker, S., & Rae, A. (2017). Everyday work investigations for safety. *Theoretical Issues in Ergonomics Science*. http://doi.org/10.1080/1463922X.2017.1356394

Håvold, J. I., Nistad, S., Skiri, A., & Ødegård, A. (2015). The human factor and simulator training for offshore anchor handling operators. *Safety Science*, *75*, 136–145. http://doi.org/10.1016/j.ssci.2015.02.001

Helmreich, R. L., Foushee, C., Benson, R., & Russini, W. (1986). Cockpit resource management: exploring the attitude-performance linkage. *Aviation Space Environment Medicine*, *57*(12), 1198–2000.

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (1999). Models of Threat, Error, and CRM in Flight Operations. In *Proceedings of the Tenth International Symposium on Aviation Psychology* (pp. 677–682).

Henderson, E. D. (1988). Air traffic controller awareness and resource training. In *Proceedings of the 7th aerospace behavioral technology conference: Operational information transfer Technology: Are we designing for the human operator* (pp. 141–144). Warrendal, PA: SAE.

Hoffman, R. R., Marx, M., Amin, R., & McDermott, P. L. (2010). Measurement for evaluating the learnability and resilience of methods of cognitive work. *Theoretical Issues in Ergonomics Science*, *11*(6), 561–575. http://doi.org/10.1080/14639220903386757

Hoffman, R. R., & Militello, L. G. (2008). *Perspectives on cognitive task analysis: Historical Origins and Modern Communities of Practice*. Boca Raton: CRC Press.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, *4*, 1–23.

Hollnagel, E. (2012). Resilience engineering and the systemic view of safety at work: Why work-as-done is not the same as work-as-imagined. In *Kongress des Gesellschaft für Arbeitswissenschaft vom 22 bis 24 Februar 2012* (pp. 19–24). Dortmund: GfA-Press.

Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Farnham: Ashgate publishing.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: foundations of cognitive systems engineering*. Boca Raton, FL, USA: CRC Press.

Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). *Resilience Engineering: Concepts and Precepts*. Farnham, UK: Ashgate.

Holt, R. W., Boehm-Davis, D. A., & Beaubien, J. M. (2001). Evaluating resource management training. In E. Salas, C. a. Bowers, & E. Eden (Eds.), *Improving teamwork in organizations: applications of resource management training* (pp. 165–188). Mahwah, NJ.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *PNAS*, *101*(46), 16385–16389.

Hopkins, A. (2007). Beyond compliance monitoring: New strategies for safety regulators. *Law and Policy*, *29*(2), 210–225. http://doi.org/10.1111/j.1467-9930.2007.00253.x

Huber, S., Wijgerden, I. Van, Witt, A. De, & Dekker, S. W. A. (2009). Learning From Organizational Incidents: Resilience Engineering for High-Risk Process Environments. *Process Safety Progress*, *28*(1), 90–95. http://doi.org/10.1002/prs

Hutchins, E. (1984). Understanding micronesian navigation. In D. Gentner & A. Stevens (Eds.), *Mental Models* (pp. 191–225). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hutchins, E. (1991). Organizing Work by Adaptation. *Organization Science*, *2*(1), 14–39. http://doi.org/10.1287/orsc.2.1.14

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA, USA: MIT Press.

Inoue, K., & Takahashi, H. (2005). Quantitative Evaluation of Consciousness Improvement in BRM Training. In *6th Annual General Assembly and Conference of the International-Association-of-Maritime-Universities* (pp. 387–395). Southampton: WIT.

Jatobá, A., Bellas, H. C., Bonfatti, R., Burns, C. M., Vidal, M. C. R., & Carvalho, P. V. R. (2016). Designing for patient risk assessment in primary health care: a case study for ergonomic work analysis. *Cognition, Technology, and Work*, *18*, 215–231. http://doi.org/10.1007/s10111-015-0355-x

Kahneman, D., & Klein, G. A. (2009). Conditions for intuitive expertise: a failure to disagree. *The American Psychologist*, *64*(6), 515–526. http://doi.org/10.1037/a0016755

Kanki, B., Helmreich, R., & Anca, J. (Eds.). (2010). *Crew Resource Management*. *Crew Resource Management*. Amsterdam, The Netherlands: Elsevier.

http://doi.org/10.1016/B978-0-12-374946-8.10014-7

Kim, C.-J. (2002). Maritime Education and Training (MET) by ship handling simulator. *Journal of the Korea Society for Simulation*, *11*(4), 81–90.

Kim, S. K., & Byun, S. N. (2011). Effects of Crew Resource Management Training on the Team Performance of Operators in an Advanced Nuclear Power Plant. *Journal of Nuclear Science and Technology*, *48*, 1256–1264. http://doi.org/10.1080/18811248.2011.9711814

Kim, S. K., Park, J. Y., & Byun, S. N. (2009). Crew Resource Management Training for Improving Team Performance of Operators in Korean Advanced Nuclear Power Plant. In *Proceedings of the 2009 IEEE IEEM* (pp. 2055–2059).

Kirkpatrick, J. D. (2016). *Kirkpatrick's Four Levels of Training Evaluation*. Alexandria, VA: ATD Press.

Klein, G. A. (1998). *Sources of Power: How People Make Decisions*. Cambridge, MA, USA: MIT Press.

Klein, G. A., Feltovich, P. J., Bradshaw, J. M., & Woods, D. D. (2005). Common Ground and coordination in joint activity. In W. B. Rouse & K. R. Boff (Eds.), *Organizational simulation* (pp. 1–42). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Klein, G. A., & Jarosz, A. (2011). A Naturalistic Study of Insight, *5*(4), 335–351. http://doi.org/10.1177/1555343411427013.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision Making in Action*. Norwood, NJ: Ablex publishing corporation. http://doi.org/10.1002/acp.2350080609

Klein, G. A., Pliske, R., Crandall, B., & Woods, D. D. (2005). Problem detection. *Cognition, Technology and Work*, *7*, 14–28. http://doi.org/10.1007/s10111-004-0166-y

Klein, G. A., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, *19*, 91–95. http://doi.org/10.1109/MIS.2004.74

Klein, K. J., & Kozlowski, S. W. J. (2000). From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research. *Organizational*

Research Methods, *3*(3), 211–236. http://doi.org/10.1177/109442810033001

Kolar, K. (2011). Resilience: Revisiting the Concept and its Utility for Social Research. *Int J Ment Health Addiction*, *9*(4), 421–433. http://doi.org/10.1007/s11469-011-9329-2

Kontogiannis, T., & Malakis, S. (2012). Remaining safe by working at the edge of compliance and adaptation: reflective practices in aviation and air traffic control. *Theoretical Issues in Ergonomics Science*, *14*(6), 565–591. http://doi.org/10.1080/1463922X.2012.672597

Krause, T. R. (1997). *The behavior-Based Safety Process: Managing Involvement for an Injury-Free Culture* (2nd ed.). New York, NY, USA: Van Nostrand Reinhold.

La Porte, T. R. (1996). High Reliability Organizations: Unlikely, Demanding and At Risk. *Journal of Contingencies & Crisis Management*, *4*(2), 60–71. http://doi.org/10.1111/j.1468-5973.1996.tb00078.x

La Porte, T. R., & Consolini, P. M. (1991). Working in Practice But Not in Theory: Theoretical Challenges of "High-reliability organizations." *Journal of Public Administration Research and Theory*, *1*(1), 19–48. Retrieved from http://www.jstor.org/stable/1181764

Lauber, J. (1984). Resource management in the cockpit. *Air Line Pilot*, *53*, 20–35.

Le Coze, J.-C. (2015). Vive la diversité! High Reliability Organisation (HRO) and Resilience Engineering (RE). *Safety Science*. http://doi.org/10.1016/j.ssci.2016.04.006

Le Coze, J.-C., & Pettersen, K. (2008). Is resilience engineering realist or constructivist? In E. Hollnagel, F. Pieri, & E. Riguad (Eds.), *3rd Symposium on Resilience Engineering* (pp. 175–184). Juan-les-Pins, France: Ecole des Mines de Paris.

Lee, E.-B., Yun, J.-H., & Jeong, T.-G. (2010). A Cyber-Training & Education Model for Tug-barge Operators. *Journal of Navigation and Port Research International Edition*, *34*(4), 287–292.

Lee, J. D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270.

http://doi.org/10.1080/00140139208967392
Lengnick-Hall, C. A., & Beck, T. E. (2008). Resilience Capacity and Strategic Agility: Prerequisites for Thriving in a Dynamic Environment. In C. P. Nemeth, E. Hollnagel, & S. W. A. Dekker (Eds.), *Resilience engineering perspectives volume 2: Preparation and restoration*. Farnham, UK: CRC Press.
Leplat, J., & Hoc, J.-M. (1983). Tache et Activite dans l'Analyse Psyholoique des situations. *Cahiers de Psychologie Cognitive*, *3*(1), 49–63.
Leveson, N. G. (2011). *Engineering a Safer World: Systems Thinking Applied to Safety*. (J. Moses, R. Neufville de, M. Heitor, G. Morgan, E. Paté-Cornell, & W. Rouse, Eds.). Cambridge, MA, USA: The MIT Press.
Lowrance, W. W. (1976). *Of Acceptable Risk: Sciene and the Determination of Safety*. Los Altos, CA, USA: William Kaufmann.
Luchins, A. S. (1942). Mechanization in Problem Solving: The Effect of Einstellung. *Psychological Monographs*, *54*(6).
Lundberg, J., & Rankin, A. (2014). Resilience and vulnerability of small flexible crisis response teams : implications for training and preparation. *Cognition, Technology & Work*, *16*, 143–155. http://doi.org/10.1007/s10111-013-0253-z
Lundberg, J., Rollenhagen, C., & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals. *Safety Science*, *47*(10), 1297–1311. http://doi.org/10.1016/j.ssci.2009.01.004
Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quartely Journal of Experimental Psychology*, *1*(1), 6–21. http://doi.org/10.1080/17470214808416738
March, J. G., & Simon, H. A. (1958). *Organizations*. Oxford, England: Wiley Organizations.
Maurino, D. E., & Murray, P. S. (2009). Crew resource management. In J. A. Wise, D. Hopkin, D. J. Garland, & B. Kantowitz (Eds.), *Handbook of Aviation Human Factors, Second Edition* (2nd ed., Vol. 411, pp. 10-1-20). Boca Raton, FL, USA: CRC Press. http://doi.org/10.1016/B978-0-12-374946-8.10014-7
McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience.

Journal of Personality and Social Psychology, *52*(6), 1258–1265. http://doi.org/10.1037/0022-3514.52.6.1258

Mearns, K., Whitaker, S., Flin, R. H., Gordon, R. P. E., & O'Connor, P. (2003). *Factoring the human into safety: Translating research into practice; Crew Resource Management Training for Offshore Operations; Volume 3 (of 3)*. *Research Report 061*. Aberdeen: Health & Safety Executive.

Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., … Tetlock, P. (2015). Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions. *Perspectives on Psyhological Sciece*, *10*(3), 267–281. http://doi.org/10.1177/1745691615577794

Merry, A. F., & McCall Smith, A. (2001). *Errors, medicine and the law*. Cambridge: cambridge university press.

Miller, J. E., Patterson, E. S., & Woods, D. D. (2005). Elicitation by critiquing as a cognitive task analysis methodology. *Cognition, Technology & Work*, *8*(2), 90–102. http://doi.org/10.1007/s10111-005-0023-7

Moffat, S., & Crichton, M. T. (2015). Investigating Non-technical Skills through Team Behavioral Markers in Oil and Gas Simulation-based Exercises. *Procedia Manufacturing*, *3*, 1241–1247. http://doi.org/10.1016/j.promfg.2015.07.259

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA GROUP. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(9), 1–6. http://doi.org/10.3736/jcim20090918

Moray, N. (2003). Monitoring , complacency , scepticism and eutactic behaviour. *International Journal of Industrial Ergonomics*, *31*, 175–178.

Moray, N., & Haudegond, S. (1998). An Absence of Vigilance Decrement in a Complex Dynamic Task. In *Proceedings of the Human Factors and Ergonomics society 42nd annual meeting* (pp. 234–238). Los Angeles, CA, USA: Sage Publications.

Moray, N., & Inagaki, T. (2000). Attention and complacency. *Theoretical Issues in Ergonomics Science*, *1*(4), 354–365. http://doi.org/10.1080/14639220052399159

Morel, G., Amalberti, R., Chauvin, C., Mörel, G., Amalberti, R., & Chauvin, C. (2008). Articulating the differences between safety and resilience: the decision-making process of professional sea-fishing skippers. *Human Factors*, *50*(1), 1–16. http://doi.org/10.1518/001872008X250683

Muir, B. M., & Moray, N. (1996). Trust in Automation II. Eperimental studies of trust and human intervention in a process control simulation. *Ergonomics*, *3*, 429–460.

Nelson, R. R., & Winter, S. G. (1985). *An evolutionary Theory of Economic Change*. Belknap Press.

Nemeth, C. P., Cook, R. I., & Woods, D. D. (2004). The messy details: Insights from the study of technical work in healthcare. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, *34*(6), 689–692. http://doi.org/10.1109/TSMCA.2004.836802

Nemeth, C. P., & Klein, G. A. (2010). The Naturalistic Decision Making Perspective. In *Wiley Encyclopedia of Operations Research and Management Science*. http://doi.org/10.1002/9780470400531.eorms0410

Norman, D. A. (1986). New views of information processing: implications for intelligent decision support systems. In E. Hollnagel, G. Mancini, & D. D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 123–136). Berlin, DE: Springer.

Norris, F. H., Stevens, S. P., Pfefferbaum, B., Wyche, K. F., & Pfefferbaum, R. L. (2008). Community Resilience as a Metaphor, Theory, Set of Capacities, and Strategy for Disaster Readiness. *Am J Community Psychol*, *41*(1–2), 127–150. http://doi.org/10.1007/s10464-007-9156-6

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. A. (2008). Crew Resource Management Training Effectiveness: A Meta-Analysis and Some Critical Needs. *The International Journal of Aviation Psychology*, *18*(4), 353–368. http://doi.org/10.1080/10508410802347044

O'Connor, P., & Flin, R. H. (2003). Crew resource management training for offshore oil production teams. *Safety Science*, *41*, 591–609. Retrieved from http://www.sciencedirect.com/science/article/pii/S0925753502000139

O'Connor, P., Flin, R. H., & Flechter, G. (2002). Methods used to evaluate the effectiveness of CRM training: A literature review. *Journal of Human Factors and Aerospace Safety*, *2*(3), 217–234.

O'Connor, P., Hahn, R., & Salas, E. (2010). The U.S. Navy's Crew Resource Management Program: The Past Present, and Recommendations For the Future. In P. O'Connor & J. Cohn (Eds.), *Human Performance Enhancements in High-Risk Environments: Insights Developments, and Future Directions from Military Research* (pp. 90–105). Santa Barbara, CA: ABC-Clio.

O'Connor, P., O'Dea, A., Flin, R. H., & Belton, S. (2008). Identifying the team skills required by nuclear power plant operations personnel. *International Journal of Industrial Ergonomics*, *38*, 1028–1037. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169814108000346

O'Dea, A., O'Connor, P., & Keogh, I. (2014). A meta-analysis of the effectiveness of crew resource management training in acute care domains. *Postgraduate Medical Journal*, *90*, 699–708. http://doi.org/10.1136/postgradmedj-2014-132800

Orasanu, J. M., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision Making in Action: Models and Methods* (pp. 3–20). Norwood, NJ: Ablex.

Ostrom, E. (1999). Coping with tragedies of the commons. *Annual Review Political Science*, *2*, 493–535.

Owen, C., Healey, A. N., & Benn, J. (2013). Widening the scope of human factors safety assessment for decommissioning. *Cognition, Technology and Work*, *15*(1), 59–66. http://doi.org/10.1007/s10111-012-0219-6

Park, S. A., Sohn, Y. W., & Kim, K. (2006). Shared Perception of Risks in Emergency Situations between Pilots and Air Traffic Controllers. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, pp. 497–500). http://doi.org/10.1177/154193120605000360

Patrick, J., & Haines, B. (1988). Training and transfer of fault-finding skill. *Ergonomics*, *31*(2), 193–210. http://doi.org/10.1080/00140138808966661

Patterson, E. S., Woods, D. D., Cook, R. I., & Render, M. L. (2007). Collaborative cross-checking to enhance resilience. *Cognition, Technology and Work*, *9*(3), 155–162. http://doi.org/10.1007/s10111-006-0054-8

Pekcan, C., Gatfield, D., & Barnett, M. L. (2005). Content and context: understanding the complexities of human behaviour in ship operation. In *RINA, Royal Institution of Naval Architects International Conference - Human Factors in Ship Design, Safety and Operation* (pp. 99–107). London, UK: Royal Institution of Naval Architects.

Pentland, B. T., & Hærem, T. (2015). Organizational Routines as Patterns of Action: Implications for Organizational Behavior. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 465–487. http://doi.org/10.1146/annurev-orgpsych-032414-111412

Pentland, B. T., & Rueter, H. H. (1994). Organizational Routines as Grammars of Action. *Administrative Science Quarterly*, *39*(3), 484–510. http://doi.org/10.2307/2393300

Perrow, C. (1984). Normal Accidents: Living With High-Risk Technologies. Basic books.

Powers, W. T. (1973). *Behavior: The control of perception*. London, UK: Wildwood House Ltd.

Rae, A., & Alexander, R. (2017). Forecasts or fortune-telling : hen are expert judgements of safety risk valid ? *Safety Science*.

Rankin, A., Dahlback, N., & Lundberg, J. (2013). A case study of factor influencing role improvisation in crisis response teams. *Cognition, Technology & Work*, *15*, 79–93. http://doi.org/10.1007/s10111-011-0186-3

Rasker, P. C., Post, W. M., & Schraagen, J. M. (2000). Effects of two types of intra-team feedback on developing a shared mental model in Command & Control teams. *Ergonomics*, *43*(8), 1167–1189. http://doi.org/10.1080/00140130050084932

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. In *IEEE Transactions on systems, man, and cybernetics* (pp. 257–266). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6313160

Rasmussen, J. (1990). Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *327*(1241), 449-460-462.

http://doi.org/10.1098/rstb.1990.0088

Rasmussen, J., & Jensen, A. (1974). Mental procedures in real-life tasks: a case study of electronic trouble shooting. *Ergonomics*, *17*(3), 293–307. http://doi.org/10.1080/00140137408931355

Reason, J. (1990). *Human Error*. Cambridge, MA: Cambridge university press.

Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Reason, J., Parker, D., & Lawton, R. (1998). Organizational controls and safety: The varieties of rule-related behavior. *Journal of Occupational and Organizational Psychology*, *71*, 289–304.

Reynaud, B. (2005). The void at the heart of rules: Routines in the context of rule-following. The case of the Paris Metro Workshop. *Industrial and Corporate Change*, *14*(5), 847–871. http://doi.org/10.1093/icc/dth073

Righi, A. W., Huber, G. J., Gomes, J. O., & de Carvalho, P. V. R. (2016). Resilience in Firefighting Emergency Response: Standardization and Resilience in Complex Systems. *IFAC-PapersOnLine*, *49*(32), 119–123. http://doi.org/10.1016/j.ifacol.2016.12.200

Roberts, K. H. (1989). New challenges in organizational research: high reliability organizations. *Industrial Crisis Quarterly*, *3*, 111–125.

Roberts, K. H. (1990). Some Characteristics of One Type of High Reliability Organization. *Organization Science*, *1*(2), 160–176. http://doi.org/10.1287/orsc.1.2.160

Rochlin, G. I. (1999). Safe operation as a social construct. *Ergonomics*, *42*(11), 1549–1560. http://doi.org/10.1080/001401399184884

Rochlin, G. I., La Porte, T. R., & Roberts, K. H. (1987). The Self-Designing High-Reliability Organization : Aircraft Carrier Flight Operations at Sea. *Naval War College Review*, *40*(4), 75–90. Retrieved from http://www.caso-db.uvek.admin.ch/Documents/MTh_SM_22.pdf%5Cnpapers3://publication/uuid/62B6CC6F-7E71-4540-94CB-32E1575D24CA

Rodrigues, M. A., Arezes, P. M., & Leão, C. P. (2015). Safety climate and its relationship with furniture companies ' safety performance and workers ' risk acceptance. *Theoretical Issues in Ergonomics Science*, *16*(4), 412–428.

http://doi.org/10.1080/1463922X.2014.1003991
Röttger, S., Vetter, S., & Kowalski, J. T. (2013). Ship management attitudes and their relation to behavior and performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(3), 659–671. http://doi.org/10.1177/0018720812461271
Runco, M. A., Dow, G., & Smith, W. (2006). Information, Experience, and Divergent Thinking: An Empirical Test. *Creativity Research Journal*, *18*(3), 269–277.
Rydenfält, C., Johansson, G., Odenrick, P., Åkerman, K., & Larsson, P. A. (2013). Compliance with the WHO surgical safety checklist: Deviations and possible improvements. *International Journal for Quality in Health Care*, *25*(2), 182–187. http://doi.org/10.1093/intqhc/mzt004
Saatcioglu, O., Asyali, E., & Cerit, A. (2005). Cooperative Learning and Teamwork Effectiveness: Impacts of Education Period on Cadets. In D. Nielson (Ed.), *International Association of Maritime Universities (IAMU) 6th Annual General Assembly and Conference.* (pp. 377–386). WIT Press. Retrieved from http://trid.trb.org/view.aspx?id=776115
Salas, E. (2005). Is there a "Big Five" in Teamwork? *Small Group Research*, *36*(5), 555–599. http://doi.org/10.1177/1046496405277134
Salas, E., Bowers, C. a., & Cannon-Bowers, J. a. (1995). Military Team Research: 10 Years of Progress. *Military Psychology*, *7*(2), 55–75. http://doi.org/10.1207/s15327876mp0702_2
Salas, E., & Burke, C. S. (2001). Team training in the skies: does crew resource management (CRM) training work? *Human Factors*, *43*(4), 641–674. Retrieved from http://hfs.sagepub.com/content/43/4/641.short
Salas, E., Diazgranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does Team Training Improve Team Performance? A Meta-Analysis. *Human Factors*, *50*(6), 903–933. http://doi.org/10.1518/001872008X375009.
Salas, E., Prince, C., Bowers, C. A., Stout, R. J., Oser, R. L., & Cannon-bowers, J. A. (1999). A Methodology for Enhancing Crew Resource Management Training. *Human Factors*, *41*(1), 161–172.

Salas, E., Wilson, K. A., Burke, C. S., & Wightman, D. C. (2006). Does crew resource management training work? An update, an extension, and some critical needs. *Human Factors*, *48*(2), 392–412. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16884057

Salas, E., Wilson, K. A., Burke, C. S., Wightman, D. C., & Howse, W. R. (2006). Crew Resource Management Training Research, Practice, and Lessons Learned. *Reviews of Human Factors and Ergonomics*, *2*(1), 35–73. http://doi.org/10.1177/1557234X0600200103

Sanda, M.-A., Johansson, J., Johansson, B., & Abrahamsson, L. (2014). Using systemic structural activity approach in identifying strategies enhancing human performance in mining production drilling activity. *Theoretical Issues in Ergonomics Science*, *15*(3), 262–282. http://doi.org/10.1080/1463922X.2012.705916

Sarter, N. B., Woods, D. D., & Billing, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd ed., pp. 1–25). Wiley.

Schaafstal, A., Schraagen, J. M., & van Berlo, M. (2000). Cognitive task analysis and innovation of training: the case of structured troubleshooting. *Human Factors*, *42*(1), 75–86. http://doi.org/10.1518/001872000779656570

Schraagen, J. M. (2011). Dealing with unforeseen complexity in the OR: the role of heedful interrelating in medical teams. *Theoretical Issues in Ergonomics Science*, *12*(3), 256–272. http://doi.org/10.1080/1464536X.2011.564481

Sheridan, T. B. (2001). Rumination on automation, 1998. *Annual Reviews in Control*, *25*, 89–97. http://doi.org/10.1016/S1367-5788(02)00009-8

Skinner, B. F. (1938). The Behavior of Organisms: An experimental analysis. *The Psychological Record*, 486. http://doi.org/10.1037/h0052216

Smith-Jentsch, K. A., Baker, D. P., Salas, E., & Cannon-Bowers, J. a. (2001). Uncovering Differences in Team Competency Requirements: The Case of Air Traffic Control Teams. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: applications of resource management training* (pp. 31–54). Mahwah, NJ: Lawrence Erlbaum Associates.

Snook, S. A. (2000). *Friendly Fire: The Accidental Shootdown of U.S. Black Hawks*

over Northern Iraq. Princeton, NJ: Princeton University Press.
Spear, S. J., & Schmidhofer, M. (2005). Ambiguity and workarounds as contributors to medical error. *Annals of Internal Medicine*, *142*(8), 627–630.
Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). *Cognitive flexbility theory: advanced knowledge acquisition in ill-structured domains*. Cambridge, MA.
Staszewski, J. J., & Davison, A. D. (2000). Mine detection training based on expert skill. In A. C. Dubey, J. F. Harvey, J. T. Broach, & R. E. Dugan (Eds.), *Proc. SPIE 4038, Detection and remediation technologies for mines and mine-like targets V* (pp. 90–101). http://doi.org/10.1117/12.396203
Szulanski, G. (2001). Appropriability and the Challenge of Scope: Banc One Routinizes Replication. In G. Dosi, R. R. Nelson, & S. Winter (Eds.), *The Nature and Dynamics of Organizational Capabilities* (pp. 69–98). Oxford University Press. http://doi.org/10.1093/0199248540.003.0004
Taylor, F. W. (1911). *The principles of scientific management*. New York, NY, USA: Harper & Brothers. Retrieved from http://books.google.co.uk/books?id=tw-YGBovG8kC
Thi, K. L. T., Jeong, J.-S., & Jeong, J.-Y. (2012). A Study on the Effective Teaching Method for Bridge Resource Management in Korea. *Journal of the Korean Society of Marine Envrionment & Safety*, *18*(6), 563–568.
Thorogood, J. L., & Crichton, M. T. (2014). Threat-and-Error Management : The Connection Between Process Safety and Practical Action at the Worksite. In *IADC/SPE Drilling conference and exhibition* (pp. 465–472).
Turner, S. F., & Fern, M. J. (2012). Examining the Stability and Variability of Routine Performances: The Effects of Experience and Context Change. *Journal of Management Studies*, *49*(8), 1407–1434. http://doi.org/10.1111/j.1467-6486.2012.01061.x
Turner, S. F., & Rindova, V. (2012). A Balancing Act: How Organizations Pursue Consistency in Routine Functioning in the Face of Ongoing Change. *Organization Science*, *23*(1), 24–46. http://doi.org/10.1287/orsc.1110.0653
Tveiten, C. K., Albrechtsen, E., Wærø, I., & Marit, A. (2012). Building resilience into emergency management. *Safety Science*, *50*(10), 1960–1966.

http://doi.org/10.1016/j.ssci.2012.03.001
Tyhurst, J. S. (1951). Individual Reactions To Community Disaster. *American Journal of Psychiatry*, *107*(10), 764–769. http://doi.org/10.1176/ajp.107.10.764
Vaughan, D. (1996). *The challenger launch decision: risky technology, culture, and deviance at NASA*. Chicago, US: The university of Chicago Press.
Vaughan, D. (1999). THE DARK SIDE OF ORGANIZATIONS: Mistake, Misconduct, and Disaster. *Annual Review of Sociology*, *25*(1), 271–305. http://doi.org/10.1146/annurev.soc.25.1.271
Vicente, K. J. (1999). *Cognitive work analysis*. London: Lawrence Erlbaum Associates.
Vicente, K. J., Mumaw, R. J., & Roth, E. M. (2004). Operator monitoring in a complex dynamic work environment: a qualitative cognitive model based on field observations. *Theoretical Issues in Ergonomics Science*, *5*(5), 359–384. http://doi.org/10.1080/14039220412331298929
Walker, B., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). Resilience, Adaptability and Transformability in Social – ecological Systems. *Ecology and Society*, *9*(2), 5. http://doi.org/10.1103/PhysRevLett.95.258101
Walker, J., & Cooper, M. (2011). Genealogies of resilience: From systems ecology of the political economy of crisis adaptation. *Security Dialogue*, *42*(2), 143–166.
Wang, A. M., Yuan, R., Li, H., Xie, W., Bakker, J. S., & Pinkster, J. (2014). Virtual simulation program and its application for challenging floatover installation of Liwan 3-1 mega topsides in South China sea. In *Proceedings of the International Offshore and Polar Engineering Conference* (Vol. 3, pp. 700–711). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84906905325&partnerID=40&md5=4782fc3184819f5827691472c5e9c9df
Watts, J. C., Woods, D. D., & Patterson, E. S. (1996). Functionally distributed coordination during anomaly response inspace shuttle mission control in Space Shuttle Mission Control. In *Human Interaction with Complex Systems '96* (pp. 68–75). Dayton, OH. http://doi.org/10.1109/HUICS.1996.549494

Weaver, S. J., Lyons, R., DiazGranados, D., Rosen, M. A., Salas, E., Oglesby, J., … King, H. B. (2010). The anatomy of health care team training and the state of practice: a critical review. *Academic Medicine : Journal of the Association of American Medical Colleges*, *85*(11), 1746–60. http://doi.org/10.1097/ACM.0b013e3181f2e907

Weick, K. E. (1987). Organizational culture as a source of high-reliability. *California Management Review*, *29*(2), 112–127.

Weick, K. E. (1990). The vulnerable system: An analysis of the Tenerife air disaster. *Journal of Management*, *16*(3), 571–593. Retrieved from http://jom.sagepub.com/content/16/3/571.short

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, *38*(3), 357–381. http://doi.org/10.2307/2393372

Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the Unexpected: Resilient Performance in an Age of Uncertainty* (2nd ed.). San Francisco: John Wiley & Sons.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (1999). Organizing for High Reliability. *Research in Organizational Behavior*, *21*, 81–123. http://doi.org/10.4135/9781452229805.n175

Westrenen, F. Van. (2014). Modelling arrival control in a vessel traffic management system. *Cognition, Technology & Work*, *16*, 501–508. http://doi.org/10.1007/s10111-014-0279-x

Woldring, M. (1999). Team Resource Management in European Air Traffic Control. *Air & Spae Europe*, *1*(1), 81–84.

Woldring, M., & Isaac, A. (1999). *Team Resource Management Test and Evaluation*. Brussels, Belgium.

Woods, D. D. (2003). Creating foresight: How resilience engineering can transform NASA’s approach to risky decision making. *US Senate Testimony of the Committee on Commerce, Science and Transportation*. Washington, DC: John McCain, chair. Retrieved from file:///Users/johnlee/Documents/Dropbox/Papers2/Files/Woods Creating foresight How resilience engineering can transform NASA?s approach to risky decision making.pdf%5Cnuuid/91085FAF-A048-423B-B021-4AEC10B1EF03

Woods, D. D., & Branlat, M. (2011). Basic Patterns in How Adaptive Systems Fail. In E. Hollnagel, J. Pariès, D. D. Woods, & J. Wreathall (Eds.), *Resilience Engineering in Practice: A Guidebook* (pp. 127–144). Farnham, England: Ashgate.

Woods, D. D., & Dekker, S. W. A. (2000). Anticipating the effects of technological change: A new era of dynamics for human factors. *Theoretical Issues in Ergonomics Science*, *1*(3), 272–282. http://doi.org/10.1080/14639220110037452

Woods, D. D., Dekker, S. W. A., Cook, R. I., Johannesen, L., & Sarter, N. (2010). *Behind human error* (2nd ed.). Farnham, UK: Ashgate Publishing limited.

Woods, D. D., & Wreathall, J. (2008). Stress-strain plots as a basis for assessing system resilience. In *Resilience engineering*.

Wright, P. (2003). Analysis of procedure following as concerned work. In E. Hollnagel (Ed.), *Handbook of Cognitive Task Design* (pp. 679–700). Mahwah, NJ: Lawrence Erlbaum Associates.

Wrigstad, J., Bergström, J., & Gustafson, P. (2017). One event, three investigations: The reproduction of a safety norm. *Safety Science*, *96*, 75–83. http://doi.org/10.1016/j.ssci.2017.03.009

Wu, Y., Miwa, T., Shimamoto, K., & Uchida, M. (2015). Development of quantitative team performance evaluation method for ERM. *WMU Journal of Maritime Affairs*, *14*(2), 333–347. http://doi.org/10.1007/s13437-015-0088-y

Yi, S., Becker, M. C., & Yi, S. (2016). Inertia in Routines: A Hidden Source of Organizational Variation. *Organization Science*, *27*(3), 782–800.

Appendix A

Characteristics of studies meeting selection criteria for conceptualisation review of chapter 2

Maritime (MAR) studies

Study	Goals	Level of effect	Compliance	Errors
(Fonne & Fredriksen, 1995)	S, P	I	C	T
(Dijkhuizen, Butter, & Koning, 1996)	S	I	n/s	H
(Byrdorf, 1998)	S, P	I	C	H
(Habberley et al., 2001)	S, P	I	B	H, T
(Barnett, Gatfield, & Pekcan, 2003)	S, P	I, T	B	H
(C.-J. Kim, 2002)	S, P	I	C	H
(Brun et al., 2005)	P	I, T	B	
(Inoue & Takahashi, 2005)	S, P	I	n/s	H
(Pekcan, Gatfield, & Barnett, 2005)	S, P	I	B	H, T
(Saatcioglu et al., 2005)	S, P	I	n/s	E
(Barsan & Hanzu-Pazara, 2010)	S	I	n/s	
(E.-B. Lee, Yun, & Jeong, 2010)	S	I	C	H
(Thi, Jeong, & Jeong, 2012)	S, P	I	n/s	H
(Dunham & Lützhöft, 2014)	S, P	I	n/s	
(Wang et al., 2014)	S, P	I, T, O	C	H
(Håvold et al., 2015)	S, P	I	B	H
(Wu et al., 2015)	S, P	I	n/s	H
(Deboo, n.d.)	S, P	I	C	H, T

Nuclear Power Industries (NPI) studies

Study	Goals	Level of effect	Compliance	Errors
(Harrington & Kello, 1991)	S, P	I	n/s	
(Harrington & Kello, 1992)	S, P	I	n/s	
(Gaddy & Wachtel, 1992)	S, P	I	n/s	T
(Davis, 2001)	S, P	I, O	C	H
(Crichton & Flin, 2004)	S, P	I	B	H, T
(O'Connor, O'Dea, Flin, & Belton, 2008)	S, P	I	C	H, T
(S. K. Kim, Park, & Byun, 2009)	S, P	I, O	B	H, T, Tn
(S. K. Kim & Byun, 2011)	S, P	I, O	B	T, Tn

Oil and Gas Industry (O&G) studies

Author	Goals	Level of effect	Compliance	Errors
(Grinde, 1994)	S	I, T, O	n/s	
(Flin, 1995)	S, P	I	n/s	E
(Flin, O'Connor, Mearns, & Gordon, 1999)	S, P	I	n/s	H, T
(Flin & O'Connor, 2001)	S, P	I	n/s	H, T, Tn
(Mearns et al., 2003)	S, P	I, O	n/s	H, T, Tn
(O'Connor & Flin, 2003)	S, P	I, O	n/s	H, T, Tn
(Crichton, 2009)	S	I, O	n/s	E
(Bailey, 2014)	S	I, T, O	n/s	
(Thorogood & Crichton, 2014)	S	I, O	C	H
(Moffat & Crichton, 2015)	S	I	B	H, T

Air Traffic Control (ATC)

Author	Goals	Level of effect	Compliance	Errors
(Henderson, 1988)	S, P	I	C	H, T
(Woldring, 1999)	S, P	I	n/s	T
(Barbarino & Isaac, 2000)	S, P	I, O	n/s	T
(Woldring & Isaac, 1999)	S, P	I, O	n/s	T
(Smith-Jentsch et al., 2001)	S, P	I, T	n/s	H
(Park, Sohn, & Kim, 2006)	S, P	I	n/s	H

Key:

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

Appendix B

CRM program module analysis

Study	Industry	Modules					
		DM	SA	LM	CP	PR	Other
Dijkhuizen et al. (1996)	MAR	0	0	3.5	0.5	0	2
Byrdorf (1998)	MAR	0	0	3	2	1	0
C.-J. Kim (2002)	MAR	0	1	3.5	3	0.5	0
Brun et al. (2005)	MAR	2	1	1	1	1	1
Pekcan et al. (2005)	MAR	0	1	1	2	0	3
Saatcioglu et al. (2005)	MAR	0	1	3	2	0	3
E.-B. Lee et al. (2010)	MAR	1	0	2	2	0	4
Thi et al. (2012)	MAR	0.5	0	1.5	2	1	0
Håvold et al. (2015)	MAR	0	0	2	2	0	4
Wu et al. (2015)	MAR	0	0	3	1	1	2
Deboo (n.d.)	MAR	1	0	7	2	0	4
Davis (2001)	NPI	2	1	1	2	0	2
S. K. Kim & Byun (2011) S. K. Kim et al. (2009)	NPI	1	1	1	1	0	1
Grinde (1994)	O&G	0	0	0	0	1	1
Flin (1995) Flin et al. (1999) Mearns et al. (2003)	O&G	1	0	1	1	1	0
Flin (1995) Flin et al. (1999) Mearns et al. (2003)	O&G	1	0	2	3	1	0
Mearns et al. (2003) O'Connor & Flin (2003)	O&G	1	1	1	1	2	0
Moffat & Crichton (2015)	O&G	1	1	1	2	1	0
Henderson (1988)	ATC	4	1	3	2	1	4
Woldring & Isaac (1999)	ATC	1	1	1	2	1	1

Key:
DM = decision making; SA = situation awareness; LM = leadership & management; CP = cooperation; PR = personal resources