The Bogeyman Unveiled

Safety and effectiveness within the Royal Netherlands Air Force

Boskeljon-Horst, L.

DOI
10.4233/uuid:315bdb50-a76b-4ce6-aa2b-fab05aa679b3

Publication date
2023

Document Version
Final published version

Citation (APA)
Boskeljon-Horst, L. (2023). *The Bogeyman Unveiled: Safety and effectiveness within the Royal Netherlands Air Force*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:315bdb50-a76b-4ce6-aa2b-fab05aa679b3

The Bogeyman Unveiled

Safety and Effectiveness within the Royal Netherlands Air Force

Leonie Boskeljon-Horst

Typography: Folkert Horst

Cover design: Nienke Horst

Printed by: Van Meltvoort, Veghel

ISBN: 9789464379754

The Bogeyman Unveiled

Safety and effectiveness within the Royal Netherlands Air Force

Dissertation

for the purpose of obtaining the degree of doctor

at Delft University of Technology

by the authority of the Rector Magnificus, prof.dr.ir. T.H.J.J. van der Hagen

chair of the Board for Doctorates

to be defended publicly on

Wednesday 21st of June 2023 at 17.30 o'clock

By

Leonie BOSKELJON-HORST

Doctorandus, Leiden University, the Netherlands

born in Eindhoven, the Netherlands

This dissertation has been approved by the promotor[s].

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof. dr. ir. P.H.A.J.M. van Gelder promotor	Delft University of Technology,
Prof. dr. S.W.A. Dekker promotor	Delft University of Technology,

Independent members:

Prof. dr. ir. M. Mulder	Delft University of Technology
Prof. dr. J. Groeneweg	Delft University of Technology
Prof. dr. I. Helsloot	Radboud University
Dr. M.S. de Vos	Leiden University Medical Centre
Cdre Mr. Drs. A.A.W.K. Appels	Director of Safety, Ministry of Defence

The Bogeyman Unveiled

Safety and effectiveness within the Royal Netherlands Air Force

Proefschrift

ter verkrijging van de graad van doctor aan de

Technische Universiteit Delft

op gezag van de Rector Magnificus prof.dr.ir. T.H.J.J. van der Hagen

voorzitter van het College voor Promoties

in het openbaar te verdedigen op

woensdag 21 juni 2023 om 17.30 uur

door

Leonie BOSKELJON-HORST

Doctorandus, Rijksuniversiteit Leiden, Nederland

geboren te Eindhoven, Nederland

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling van de promotiecommissie:

Rector Magnificus,	Voorzitter
Prof. dr. ir. P.H.A.J.M. van Gelder	Technische Universiteit Delft, promotor
Prof. dr. S.W.A. Dekker	Technische Universiteit Delft, promotor

Onafhankelijke leden:

Prof. dr. ir. M. Mulder	Delft University of Technology
Prof. dr. J. Groeneweg	Delft University of Technology
Prof. dr. I. Helsloot	Radboud University
Dr. M.S. de Vos	Leiden University Medical Centre
Cdre Mr. Drs. A.A.W.K. Appels	Director of Safety, Ministry of Defence

To Thomas and Max, my greatest accomplishments

To Richard, the love of my life

To my father, who showed me the way

Acknowledgements

I would like to thank Pieter van Gelder and Simone Sillem for their support and trust during the process of completing this thesis.

I owe a special debt of thanks to Sidney Dekker. You have been an inspiration for over 15 years, since I first read one of your articles. It has been my mission ever since to introduce your ideas to the RNLAF and change not only the way we look at safety, but also the way we look at all the men and women of the RNLAF. Thank you for your guidance, your support and valuable insights.

I owe another special debt of thanks to my father, RNLAF Retired Major General Folkert Horst, who has spent countless hours reading draft versions. Thank you for introducing me to the RNLAF and showing me what being a good RNLAF officer is like.

Also, I would like to thank the other co-authors who have significantly contributed to the studies presented in this thesis. Robert Jan de Boer, thank you for guiding me along a pretty steep learning curve and all your valuable feedback. Vincent Steinmetz, thank you for all the conversations we have had about Safety II and all your support during my writing. Eva van Baarle, thank you for your guidance and your constructive feedback. It is a privilege to work with you. Anke Snoek, thank you for your confidence in my ability to finish this thesis and good company.

Furthermore, I would like to thank my colleagues that have supported me one way or the other: Joyce Neef, Edzard Boland, Michel de Rivecourt, Ralph Tier, Arthur Rademakers, Ron Koppes and Gwendolyn Bakx.

Finally, I would like to thank the men and women of the RNLAF Defence Helicopter Command for allowing me to use their experiences, ideas, frustrations and stories for this research. I hope this research will be used to create the necessary space for you to use your expertise and to make your daily work less complex, so that you can be even more successful.

TABLE OF CONTENTS

Summary

In my nearly 24 years as an aviation psychologist in the Royal Netherlands Air Force (RNLAF), I have seen first-hand the dynamics and complexity of the daily work situations of pilots and crew. I have also seen an organisation trying to enhance safety, prevent negative incidents and advocate the importance of learning. And I have seen this same organisation failing at all three. I became convinced that giving operators more discretionary space to use their expertise, developing a better understanding of successes and focussing on restorative instead of retributive justice are the keys to enhancing the safety and the effectiveness of the RNLAF.

The RNLAF is currently transitioning to a Fifth Generation Air Force (5GAF). In order to stay relevant and gain the competitive advantage it needs, the RNLAF not only wants new weaponry but also a different management style that will foster different behaviour in employees. Specifically, the 5GAF focusses on trust, accountability, more freedom and space to employees and more room for self-organisation rather than top-down control. In my opinion, more discretionary space, a better understanding of successes and focussing on restorative justice would provide more competitive advantage than any weapons system we could acquire.

Within the RNLAF context I studied the (in)ability to make sense of retrieved safety information and observations, how safety and effectiveness is achieved, and how this is hampered by safety beliefs and retributive response to undesired outcomes. My central research question is: **How can we describe and enhance the safety and effectiveness of the RNLAF?** The sub-questions all focus on understanding aspects of safety with the intention of enhancing it:

Three key concepts are relevant to my research: safety culture, just culture and compliance vs. adaptation. These three concepts take up a significant part of the literature focused on enhancing safety and therefore, put together, they might provide a solid explanation for not only the safety an organisation achieves but also the stalemate and plateauing results an organisation meets when trying to further enhance safety. Safety culture, just culture and compliance are interrelated. Rules and procedures are regarded essential elements of a safety culture. The response to violating rules and procedures shows the just culture of an organisation.

Safety culture, just culture and compliance provide a common thread in the safety documents of the Defence organisation appearing in the past six years. Documents show there is outside pressure to enhance the safety culture in the Defence organisation, an underdeveloped restorative just culture and a recognition that both compliance and proactive intervention (adaptation) are needed.

The sub-questions are answered as follows.

Sub-question 1: What is the predictive value of a safety culture assessment regarding the accident-proneness of an operational squadron?

To study the predictive value of safety culture assessment (SCA) a comparison was made between the results of an SCA and the findings of an incident investigation. The results of this comparison showed that the incident investigation led to different results and a different safety culture description than the safety culture maturity assessment. This led to the conclusion that the SCA has limited predictive value

for the safety culture during an accident. In fact, the SCA inversely predicted the outcome. SCA factors indicative of a mature, proactive and generative, were determined by the investigation committee to be the causes of this particular accident. This means that a maturity SCA seems to provide little guidance when trying to predict safety outcomes or enhance safety and effectiveness.

Sub-question 2: What are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys?

To study the differences regarding goal realisation when applying two conflicting rule management paradigms, more than 30 scenarios with goal conflicts were collected. These goal conflicts were frequently solved locally with routines dubbed "local ingenuity". The local ingenuity was either created within the existing rule base or non-transparent to the organisation outside the unit and explained by the different paradigms regarding rule management. The RNLAF is applying a top-down classical, rational approach to rules, whereas the operators abide by a bottom-up constructivist view rules, as seen in the examples of local ingenuity. This study shows that the classical paradigm is insufficient to understand how work is done. It constrains the development and transparency of local ingenuity and results in the inhibition of goal conflict solutions. As a result local ingenuity is suboptimal as is goal realisation, leading to overall suboptimal effectiveness of the unit studied.

Sub-question 3: Why do differences between the way work is done and prescribed endure despite their undesirability?

Conflicting goals and the resulting local ingenuity lead to gaps between the way work is done and prescribed. These gaps are considered undesirable, yet attempts to breach the status quo of gaps remain stymied because of a forceful dynamic balance of interests that both produce and retain the status quo. The pressure to sustain the compliance promise made to politicians, results in unsuccessful attempts to gain discretionary space and fewer rules. This study showed both a simultaneously retaining *and* denial of noncompliance at almost all hierarchical levels studied and a belief system, reminiscent of a bogeyman, in which compliance equals safety. It is because of this belief system gaps between paper and practice endure.

Sub-question 4: What complexities arise when attempting to foster a restorative just culture in a mostly retributive organisation?

To examine the complexities of fostering a restorative just culture in practice, a case study of a safety standdown in the RNLAF was examined. The findings show that four complexities are of great importance for a restorative response: 1) a willingness to be vulnerable without being punished for this; 2) moral courage and a fair attitude, even if it means sacrificing one's own interests; 3) a management that does not punish commanding officers for responding restoratively; and 4) a desire to learn from the mistake instead of punishing for it. This study also demonstrates how hard it is to aspire to a restorative just culture in a predominantly retributive setting and culture. Examples that fostered restorative openness were threatened by the deeply engrained response to punish in the RNLAF. The four complexities need to be overcome to strengthen a restorative response and build a restorative just culture that will enhance the safety of the organization.

Sub-question 5: Are micro-experiments enough to resolve the enduring gaps between work-as-imagined and work-as-done?

It was studied if an approach combining the framework as proposed by Hale and Borys (2013b) with micro-experiments would result in a reduction of the gap between prescription and practice, a change in rules to better match daily operations and an understanding whether this method can break through the organisational complicity. This study shows that Dekker and De Boer rightly proposed micro-experiments as effective interventions and means to gain shop floor commitment, but that the difficulty of actually modifying rules has been overlooked. This was impeded by the tension between compliance and adaptation and the presence of a bogeyman belief regarding non-compliance. The proposed approach used in this study is insufficient to break through the organisational complicity. Adding a focus on subjects such as psychological safety, restorative justice, psychological contracts and betrayal and moral disengagement seems needed.

The research question is subsequently answered as follows:

Describing and enhancing both safety and effectiveness within the RNLAF requires understanding and enhancing local ingenuity and making this transparent outside the unit it is applied. This subsequently requires a change in the belief system, embracing the constructivist paradigm of rule management and a restorative just culture.

The literature in two scientific domains – on complex systems and on safety as an adaptive capacity – supports the answer to the research question. First, complex systems are dynamic, involving a large number of continuously changing interacting elements. The relationships are non-linear and the patterns unpredictable, which make adaptation necessary to respond to the emerging patterns. The local ingenuity found is the result of operators adapting to the complexity of the local contexts in which predictability, stability and control do not apply. Second, because complex systems require adaption, the bottom-up constructivist approach to rule management and the Safety II/Safety Differently approach provide a better fit than the top-down classical approach or Safety I approach. Seen from these paradigms (constructivist/Safety II), safety means the presence of adaptive capacity to make things go right in dynamic, surprising and uncertain situations. The capacity index, created by Dekker and Tooma as an alternative to Safety I measures, provides six demonstrable and measurable capacities that provide insight into the safety of an organisation of which five provide support for the research described in this thesis.

What is needed to enhance RNLAF safety and effectiveness is a paradigm shift from model 1 to model 2 thinking, not only specifically for rule management, but also more generally for to manageability and predictability. Model 1 thinking assumes the correctness of rules, predictability and manageability. Rules provide control over the concept of safety. But many processes in the RNLAF fall into the realm of complexity and complex systems. This means model 2 thinking is needed to put the compliance commitment in perspective and enhance both safety and effectiveness. Furthermore, it needs a paradigm shift from a retributive to a restorative just culture. Finally, the local ingenuity found shows that operators need more space in which to use their expertise. This requires freedom-in-a-frame or discretionary space for professional judgement. Perhaps what the RNLAF needs to enhance safety and effectiveness can best be explained by General George S. Patton: '*Never tell people how to do things. Tell them what to do, and they will surprise you with their ingenuity*.'

Samenvatting

In mijn bijna 24 jaar als vliegerpsycholoog bij de Koninklijke Luchtmacht (KLu) heb ik de dynamiek en complexiteit van de dagelijkse werksituaties van vliegers en bemanningsleden aanschouwd. Ik heb een KLu gezien dat probeert de veiligheid te vergroten, incidenten te voorkomen en te leren. En ik heb diezelfde KLu organisatie zien falen op alle drie de gebieden. Ik ben ervan overtuigd geraakt dat het geven van meer discretionaire ruimte aan operators om hun expertise te gebruiken, het ontwikkelen van een beter begrip van successen en het focussen op herstellende in plaats van bestraffende rechtvaardigheid de sleutels zijn tot het verbeteren van de veiligheid en de effectiviteit van de KLu.

De KLu is momenteel bezig met een transitie naar een vijfde generatie luchtmacht (5GAF). Om relevant te blijven en een militair overwicht te behalen, wil de KLu niet alleen nieuwe wapens, maar ook een andere managementstijl die ander gedrag bij medewerkers zal stimuleren. In het bijzonder richt de 5GAF zich op vertrouwen, verantwoording, meer vrijheid en ruimte voor werknemers en meer ruimte voor zelforganisatie in plaats van top-down controle. Naar mijn mening zouden meer discretionaire ruimte, een beter begrip van successen en aandacht voor herstelrecht meer militair overwicht opleveren dan welk wapensysteem dan ook.

Binnen de KLu bestudeerde ik het (on)vermogen om betekenis te geven aan veiligheidsinformatie en observaties, hoe veiligheid en effectiviteit worden bereikt, en hoe dit wordt belemmerd door veiligheidsovertuigingen en bestraffende reacties op ongewenste uitkomsten. Mijn centrale onderzoeksvraag is: **Hoe kunnen we de veiligheid en effectiviteit van de KLu beschrijven en verbeteren?** De deelvragen richten zich op het begrijpen van aspecten van veiligheid, met de bedoeling deze te vergroten:

Drie sleutelbegrippen zijn relevant voor mijn onderzoek: veiligheidscultuur, just cultuur en compliance versus aanpassing. Deze drie concepten nemen een belangrijk deel van de literatuur in beslag dat gericht is op het verbeteren van veiligheid en bieden daarom samen mogelijk een solide verklaring voor niet alleen de veiligheid die een organisatie bereikt, maar ook voor de impasse die een organisatie bereikt wanneer zij probeert de veiligheid verder te verbeteren. Veiligheidscultuur, just culture en compliance hangen met elkaar samen. Regels en procedures worden gezien als essentiële elementen van een veiligheidscultuur. De reactie op het overtreden van regels en procedures toont de just culture van een organisatie.

Veiligheidscultuur, just culture en compliance vormen een rode draad in de veiligheidsdocumenten van Defensie die in de afgelopen zes jaar zijn verschenen. Uit de documenten blijkt een druk van buitenaf om de veiligheidscultuur in de Defensieorganisatie te verbeteren, een onderontwikkelde herstelgerichte just culture en een erkenning dat zowel compliance als proactief ingrijpen (aanpassing) nodig is.

De deelvragen worden als volgt beantwoord.

Deelvraag 1: Wat is de voorspellende waarde van een veiligheidscultuur assessment met betrekking tot de ongeval-gevoeligheid van een operationeel squadron?

Om de voorspellende waarde van safety culture assessment (SCA) te onderzoeken is een vergelijking gemaakt tussen de resultaten van een SCA en de bevindingen van een incidentenonderzoek. Uit deze

vergelijking bleek dat het incidentenonderzoek tot andere resultaten en een andere beschrijving van de veiligheidscultuur leidde dan de safety culture maturity assessment. Dit leidde tot de conclusie dat de SCA een beperkte voorspellende waarde heeft voor de veiligheidscultuur tijdens een ongeval. In feite voorspelde de SCA de uitkomst omgekeerd. SCA-factoren die wijzen op een volwassen, proactieve en generatieve cultuur, werden door de onderzoekscommissie als oorzaak van dit specifieke ongeval aangewezen. Dit betekent dat een volwassen SCA weinig houvast lijkt te bieden bij pogingen om veiligheidsresultaten te voorspellen of de veiligheid en effectiviteit te verbeteren.

Deelvraag 2: Wat zijn de verschillen met betrekking tot doelrealisatie bij het toepassen van twee conflicterende paradigma's met betrekking tot regelperceptie en -management zoals beschreven door Hale & Borys?

Om de verschillen met betrekking tot doelrealisatie bij het toepassen van twee conflicterende paradigma's met betrekking tot regelgeving te bestuderen, werden meer dan 30 scenario's met doelconflicten verzameld. Deze doelconflicten werden vaak lokaal opgelost met routines die "lokale vindingrijkheid" werden genoemd. Deze lokale vindingrijkheid werd ofwel gecreëerd binnen de bestaande regelgeving ofwel niet-transparant voor de organisatie buiten de eenheid en verklaard door de verschillende paradigma's met betrekking tot regelgeving. De KLu hanteert een top-down klassieke, rationele benadering van regels, terwijl de operators een bottom-up constructivistische kijk op regels hebben, zoals blijkt uit de voorbeelden van lokale vindingrijkheid. Deze studie toont aan dat het klassieke paradigma onvoldoende is om te begrijpen hoe werk wordt gedaan. Het beperkt de ontwikkeling en transparantie van lokale vindingrijkheid en remt doelconflictoplossingen. Als gevolg daarvan is lokale vindingrijkheid en doelrealisatie suboptimaal, wat leidt tot algemene suboptimale effectiviteit van de bestudeerde eenheid.

Deelvraag 3: Waarom blijven verschillen tussen de manier waarop werk wordt gedaan en voorgeschreven bestaan ondanks hun onwenselijkheid?

Tegenstrijdige doelen en de daaruit voortvloeiende lokale vindingrijkheid leiden tot verschillen tussen de manier waarop het werk wordt gedaan en is voorgeschreven. Deze verschillen worden als ongewenst beschouwd, maar pogingen om de status-quo van deze verschillen te doorbreken worden tegengehouden door een krachtige dynamisch evenwicht van belangen die de status-quo zowel produceren als in stand houden. De druk om de aan politici gedane belofte van compliance na te komen, resulteert in mislukte pogingen om discretionaire ruimte en minder regels te krijgen. Deze studie toonde aan dat op bijna alle onderzochte hiërarchische niveaus noncompliance tegelijkertijd wordt gehandhaafd en ontkend, en dat er een overtuiging bestaat, dat doet denken aan een boeman, waarin compliance gelijkstaat aan veiligheid. Door deze overtuiging blijft er een verschil bestaan tussen papier en praktijk.

Deelvraag 4: Welke complicaties doen zich voor bij pogingen om een herstelgerichte rechtvaardige cultuur te bevorderen in een overwegend bestraffende organisatie?

Om de complexiteit van het bevorderen van een herstelgerichte just culture in de praktijk te onderzoeken, is een casestudy van een safety standdown in de KLu onderzocht. Uit de bevindingen blijkt dat vier complexiteiten van groot belang zijn voor een herstelgerichte respons: 1) de bereidheid om zich kwetsbaar op te stellen zonder daarvoor gestraft te worden; 2) morele moed en een rechtvaardige houding, ook als dat betekent dat het eigen belang moet worden opgeofferd; 3) een management dat commandanten niet straft voor een herstelgerichte reactie; en 4) de wens om te leren

van de fout in plaats van ervoor te straffen. Deze studie toont ook aan hoe moeilijk het is om een herstelgerichte cultuur na te streven in een overwegend bestraffende omgeving en cultuur. Voorbeelden die herstellende openheid bevorderden werden bedreigd door de diepgewortelde reactie om te straffen in de KLu. De vier complexiteiten moeten worden overwonnen om een herstelgerichte respons te versterken en een herstelgerichte rechtvaardige cultuur op te bouwen die de veiligheid van de organisatie zal vergroten.

Deelvraag 5: Zijn micro-experimenten voldoende om de verschillen tussen papier en praktijk te verminderen?

Onderzocht werd of een aanpak die het framework zoals voorgesteld door Hale en Borys combineert met micro-experimenten zou leiden tot een verkleining van het verschil tussen papier en praktijk, resulteert in een verandering van regels om beter aan te sluiten bij de dagelijkse werkzaamheden en of deze methode de organisatorische geheimhouding kan doorbreken. Deze studie laat zien dat Dekker en De Boer micro-experimenten terecht hebben voorgesteld als effectieve interventies en middelen om betrokkenheid van de werkvloer te krijgen, maar dat de moeilijkheid van het daadwerkelijk aanpassen van regels over het hoofd is gezien. Dit werd belemmerd door de spanning tussen compliance en aanpassing en de aanwezigheid van een boemangeloof ten aanzien van noncompliance. De in deze studie voorgestelde aanpak is onvoldoende om de organisatorische geheimhouding te doorbreken. Een focus op onderwerpen als psychologische veiligheid, herstelrecht, psychologische contracten en verraad en morele ontkoppeling lijkt nodig.

De onderzoeksvraag wordt vervolgens als volgt beantwoord:

Het beschrijven en vergroten van zowel de veiligheid als de effectiviteit binnen de KLu vraagt om het begrijpen en vergroten van lokale vindingrijkheid en het transparant maken hiervan buiten de eenheid waar het wordt toegepast. Dit vereist vervolgens een verandering in veiligheidsopvattingen, waarbij het constructivistische paradigma van regelgeving en een herstelgerichte just culture worden omarmd.

Wetenschappelijke literatuur over complexe systemen en over veiligheid als aanpassingsvermogen ondersteunt het antwoord op de onderzoeksvraag. Ten eerste zijn complexe systemen dynamisch, met een groot aantal voortdurend veranderende, op elkaar inwerkende elementen. De relaties zijn niet-lineair en de patronen onvoorspelbaar, waardoor aanpassing noodzakelijk is om op veranderingen in te spelen. De gevonden lokale vindingrijkheid is het resultaat van de aanpassing van de operators aan de complexiteit van de lokale context waarin voorspelbaarheid, stabiliteit en controle niet van toepassing zijn. Ten tweede, omdat complexe systemen aanpassing vereisen, passen de bottom-up constructivistische benadering van regelgeving en de Safety II/Safety Differently benadering beter dan de top-down klassieke benadering of Safety I benadering. Bezien vanuit deze paradigma's (constructivistisch/veiligheid II) betekent veiligheid de aanwezigheid van aanpassingsvermogen om in dynamische, verrassende en onzekere situaties de zaken goed te laten verlopen. De capaciteitsindex, gecreëerd door Dekker en Tooma als alternatief voor Safety I maatregelen, biedt zes aantoonbare en meetbare capaciteiten die inzicht geven in de veiligheid van een organisatie, waarvan er vijf ondersteuning bieden voor het in dit proefschrift beschreven onderzoek.

Wat nodig is om de veiligheid en effectiviteit van de KLu te vergroten is een paradigmaverschuiving van model 1- naar model 2-denken, niet alleen specifiek voor regelgeving, maar ook meer in het algemeen voor beheersbaarheid en voorspelbaarheid. Model 1-denken gaat uit van de juistheid van regels,

voorspelbaarheid en beheersbaarheid. Regels geven controle over het begrip veiligheid. Maar veel processen binnen de KLu vallen in het domein van complexiteit en complexe systemen. Dit betekent dat model 2-denken nodig is om de compliance belofte in perspectief te plaatsen en zowel de veiligheid als de effectiviteit te vergroten. Bovendien is een paradigmaverschuiving van een retributieve naar een restaurative just cultuur nodig. Ten slotte blijkt uit de gevonden lokale vindingrijkheid dat operators meer ruimte nodig hebben om hun expertise in te zetten. Dit vereist vrijheid-in-een-kader of discretionaire ruimte voor professioneel oordeel. Wat de KLu nodig heeft om de veiligheid en effectiviteit te vergroten kan misschien het best worden uitgelegd door generaal George S. Patton: "*Vertel mensen nooit hoe ze dingen moeten doen. Vertel ze wat ze moeten doen, en ze zullen je verrassen met hun vindingrijkheid*."

Page intentionally left blank

Chapter 1 Introduction

1. Introduction

In 2016, a squadron of the Royal Netherlands Air Force (RNLAF) participated in an international military flight exercise conducted abroad. During the flight back to the Netherlands, the Dutch aircrew was involved in three separate events.

The first event happened at the end of the exercise, when the participating nations provided a flypast to thank all the support personnel for their work during the exercise. A flypast is a low-altitude flight by one or more aircraft above a public gathering. This manoeuvre was planned and briefed but executed differently from the plan. To set up the flypast, the formation of helicopters was supposed to make a 180-degree turn right. The second-to-last helicopter miscalculated the turn, hampering the last, Dutch, helicopter, which abruptly had to reduce speed to avoid a collision. Unfortunately, and unknown to all the pilots in the flypast, the support staff had started breaking up camp the day before. This involved loosening all the tent pegs but leaving the tents in place. As the tents were no longer fixed to the ground, when the Dutch helicopter reduced speed, its downwind blew over and damaged a few tents. After leaving the exercise area, the Dutch crew became aware of the damage they had incurred. Feeling bad about what had happened, they flew back to the Netherlands.

The second event happened en route to the Netherlands. Not wanting to let these flight hours go to waste, the Dutch crew planned a route that included low-level flying and a low approach to another airfield as part of their training. Because it was a civil airfield, they needed to pay a fee to conduct the low approach but, unfortunately, they had forgotten to do this. Their unpaid low approach resulted in an angry phone call to the squadron command followed by a bill for the due fee.

About 20 minutes before arriving at their destination, the crew was involved in a third event. The military exercise had lasted three weeks during which the participants ate, slept and worked together, pretty much 24/7. Naturally a strong bond emerged. Just prior to returning home, a member of the Dutch crew learned that his wife and three kids were ill with the flu. Before the exercise his wife had dropped him off at the airbase and she would pick him up there afterwards. To get to the airbase she had to drive 90 minutes with three sick kids in the back and feeling sick herself. On the flight home the crew discussed the situation and came up with an alternative. They would do an extra landing in the low-level flying area to drop this crew member off so that his wife would only have a 30-minute drive. This crew member was known for his team spirit and cooperative nature. Dropping him off early felt the right thing to do, as it would return just one of the thousands of favours he had done for the crew. Moreover, the motto of the RNLAF is (translated) “one team, one job”, emphasising the importance of teamwork. It is against the rules to fly a helicopter with an incomplete crew. They compensated for this by putting someone else in his position. However, this person was not an official member of the crew.

After landing at the home airbase, the pilot-in-command (PIC) went straight to his commanding officer (CO) and told him about the damaged tents and low-approach fee. The minute they had landed the PIC knew they had made the wrong decision regarding the drop-off, since the rules state that it is not allowed to fly a helicopter with a crew member missing. Because the PIC already had to tell the CO about the first two events, he felt reluctant to mention the drop-off as well. Nonetheless, that evening he called the CO from home and filled him in on the drop-off situation. The PIC insisted he was solely responsible for all that had happened.

The CO assembled an investigating committee to look into the three events. After the investigation, the PIC received no fewer than five punishments. First, he was relieved of his PIC status for six months. He would remain in his leadership position but only for administrative work. Second, he had to write an apology to the nation that owned the tents. Third, he had to give mandatory briefings to his own and other squadrons explaining what he had done. Fourth, on top of his already extensive workload he had to arrange the next international exercise and, finally, a formal reprimand would stay in his file for a year.

Looking back on this day, the PIC remembers that he felt disappointed by his CO's formal response:

> *"Of course, we, or better yet, I made mistakes. I was the PIC and even though it was a group decision, I am responsible and accountable. I knew when we landed that I'd made a mistake. And I'd learned from it long before anyone had formed an opinion or said anything about it. The way I was [treated] just felt disproportionate to what had happened. It didn't seem to fit my CO's personality, at least not as I knew him. What I needed then was an effort to really understand how all of this happened and a more reasonable response.*
>
> *"I felt the weight of the consequences for an entire year. I just don't understand why the response was so harsh, especially from my own CO. The airbase commander was far more understanding. He said, "These things happen, I have to punish you[1], but I do understand." And so did the nation that owned the tents. The reply I got to my apology basically said, "What are you worried about? These things happen in exercises. We fix the tents and move on to the next exercise." So, it is rather strange how much blame I received from my own CO. […] I see myself as a rather obedient pilot and because of one incident I am made an example for my colleagues. And you know what's hurt me the most? In the goodbye speech they gave me when I resigned, all they talked about was this incident. It's as if, after all my years with the squadron, all that defines me is this incident. The same has happened to colleagues. Everything you've done is reduced to one particular outcome."*

The CO recalls:

> *"I'd just become the CO of this squadron. I was still in my first 100 days in which you try to define yourself, show what you as a CO think is important. At some point I learned about the damage and pretty soon I received questions from my own CO as to what had happened. It was suggested that it might need to be investigated. I did feel pressure from outside my own squadron, so to speak. […] What intrigued me about this occurrence was that it involved this particular PIC. Of all my pilots, he was the last one I expected to appear in front of me. […]*
>
> *"In my opinion that flypast was done a little hastily, the necessary reduction in speed caused the downwash damage. Then as they went on, they landed somewhere, forgetting to pay the landing fee, to drop off one of their crew, thinking we're nearly back, ready for the weekend. […] I wanted to have an honest talk with this crew, because as a CO I can have some influence on the consequences. I wanted to see if we could look at this from a 'just culture' perspective and learn from it. At the same time, I knew and felt that my own COs were expecting me to*

[1] The airbase commander gave the PIC the formal reprimand. His commanding officer chose the other four consequences.

investigate the events and punish the crew. […] What was important to me is that they could have made good choices – landing after they damaged the tents, apologising over the radio – but they didn't and only made the situation worse by dropping off a crew member without paying the landing fee. Nobody asked at any point if they were being smart about it. […] I really thought [the PIC] needed to be straightened out on this matter. […] If I hadn't intervened, someone else would have. […]

"The investigation offered us a way out, a way of leaving all of this behind us, closure, letting us move on, the PIC and me. So, he lost his PIC status for six months, because after that we needed him again, we couldn't afford to miss him as PIC. I think he accepted the consequences but if we did the right thing, if I did the right thing, I don't know. […]

"In this organisation, senior management quickly micromanages anything that goes wrong and might have a political impact, poking their long screwdriver right into the work floor. Which is incredibly annoying. I mean, people don't come to work intending to screw up. […]

"I've been in a situation where it was all done according to the rules, with permission, but a journalist took a picture and then I had to spend [a whole] weekend making phone calls justifying myself right up to the Minister of Defence because questions were asked about that picture. No one backs you up, so it all came back on me because I had approved this particular flight. […]

"So, we talk a lot about discretionary space, true, but only as long as we like the results. […] There is so much fear in our organisation, that the outside world will have an opinion, so we instantly assume that things are done with bad intentions. And that's when they use the long screwdriver. It spears through us all, so we look like those little dolls in table football. When you turn that screwdriver all the dolls, meaning everyone in the chain of command, spins the same way and says the same thing, usually something like "Yes, error, mistake, punish." […] We all spin with this screwdriver because we're all afraid of the negative consequences for our career if we don't."

In my nearly 24 years as an aviation psychologist in the RNLAF, I have seen first-hand the dynamics and complexity of the daily work situations of pilots and crew. I have heard many stories like the ones above, which dismiss the complexities and blame individuals for undesired outcomes. I have noticed a judgemental attitude throughout the RNLAF and the impact it has on personnel who are just trying to do the best they can, just the way they always do before and after an incident. I have seen both informal and formal punishment for unwanted outcomes. And I have witnessed how aircrews, maintenance crews, flight controllers, fire fighters and everyone else on the RNLAF work floor – summarised in the term 'operators' – are responsible for the successes of the RNLAF, often achieved without management support, as this research will reveal. I have also seen an organisation trying to enhance safety, prevent negative incidents and advocate the importance of learning. And finally, I have seen this same organisation failing at all three.

With the insights gained in my long experience as an aviation psychologist, I began pondering:

- What exactly does safety mean to the RNLAF and its operators?
- How do the RNLAF and its operators incorporate safety in daily operations?

- What is the response when it comes to [negative] incidents?
- What does compliance mean and how does it fit in with all of this?
- What do operators do that deviates from rules and procedures and is it really wrong and/or dangerous?
- What role does the organisation have in encouraging the behaviour seen in this and many other stories?
- What impact does the Defence/RNLAF view on safety, human behaviour and compliance have on the effectiveness of the RNLAF?

Round about 2006 I got to know the work of Erik Hollnagel [Senior Professor of Patient Safety at Jönköping University (Sweden)] and Sidney Dekker [a professor at Griffith University in Brisbane (Australia), where he founded the Safety Science Innovation Lab]. After attending one of Dekker's lectures where he told the audience about Eve and the serpent (Dekker, 2007d) and the impact of the 'old view' (Dekker, 2004) on an organisation, it dawned on me that taking another way of looking at incidents and humans could make all the difference. Thus began my journey of reading, learning and spreading the word in the RNLAF. After speaking up about the 'wrong way' of doing incident and accident investigation and proving that all the reports came to the same conclusions and made the same recommendations (either we as an organisation really do not learn in the RNLAF, have a chronic lack of creativity or never really look beyond the surface of a mishap), I was appointed the new Head of Occurrence Investigations in early 2018. Then my real mission began: changing the entire investigation process, entrenching the RNLAF organisation in the new 'Safety II'/'Safety Differently' vision (Dekker, 2006) (Dekker, 2014b) (Hollnagel, 2014), enhancing the just culture inside our organisation and showing how much the RNLAF depends and relies on the professionalism and expert judgement of its operators. I am convinced that giving operators more discretionary space to use their expertise, developing a better understanding of successes and focussing on restorative instead of retributive justice are the keys to enhancing the safety and the effectiveness of the RNLAF. In my opinion, these three elements would provide more competitive advantage than any weapons system the RNLAF could acquire. Furthermore, if, as a science, it is known since 1947 (Dekker, 2015) that human behaviour is inextricably linked to the surrounding context, that people do what makes sense to them given the circumstances and knowledge of the time (Dekker, 2006), it seems a moral and ethical duty to understand that context as well as how our view on human behaviour determines the effectiveness of the RNLAF.

With this dissertation I set out to do just that: in the RNLAF context I studied the (in)ability to make sense of retrieved safety information and observations, how safety and effectiveness is achieved, and how this is hampered by safety beliefs and retributive response to undesired outcomes.

My central research question is: **How can the safety and effectiveness of the RNLAF be described and enhanced?** The sub-questions all focus on understanding aspects of safety with the intention of enhancing it:

1. What is the predictive value of a safety culture assessment regarding the accident-proneness of an operational squadron?
2. What are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013)?
3. Why do differences between the way work is done and is prescribed endure despite their undesirability?

4. What complexities arise when attempting to foster a restorative just culture in a mostly retributive organisation?
5. Are micro-experiments enough to reduce the enduring gaps between work-as-imagined and work-as-done?

This research adds to the safety research domain and contributes to the transition of the RNLAF to a Fifth Generation Air Force (5GAF).

The next section (2) introduces the context of the organisation. Section 3 follows with the theoretical background of safety paradigms, safety culture, just culture and compliance. Section 4 provides the rationale for this dissertation, while section 5 explains the overall research question and sub-questions. Section 6 concludes this Introduction with an outline of the remaining chapters.

2. The context: Royal Netherlands Air Force and Defence Helicopter Command

The research described in this dissertation was conducted in the Defence Helicopter Command of the RNLAF, one of the armed-force services of the Dutch Ministry of Defence (MoD). The MoD contains seven organisational elements in total: central staff responsible for defence policy, four services responsible for mission-ready military personnel, and two support commands responsible for supportive products and services. The four services and two support commands are directed by the Chief of Defence, the highest military position in the MoD.

The RNLAF provides security from the air. Directed by the commander, administered by the central staff, the RNLAF contains four divisions: Air Combat Command, Air Mobility Command, Air Support Command and Defence Helicopter Command (DHC).

DHC is responsible for the availability of helicopter and support units. Its main mission is to provide helicopter air support for the Marines, Air Mobile Brigade and to support UN and NATO deployment missions. In practice, this means conducting reconnaissance flights, facilitating ground operations by transporting personnel and equipment and providing air protection during transport. The DHC operates at Gilze Rijen Air Force base, Deelen Military Airbase and De Kooy Navy Airbase. Gilze Rijen and De Kooy maintain four operational flight squadrons. Three of the four squadrons participated in this research; the first squadron is described in Chapter 2, the second features in Chapters 3, 4 and 6, and the findings on the third operational squadron are detailed in Chapter 5.

2.1 Fifth Generation Air Force

Working on making the transition to the Fifth Generation Air Force (5GAF), the RNLAF is focused on increasing decisiveness, agility and resilience. An RNLAF leaflet on 5GAF published in 2020 states the following:

> *The security situation in Europe and around the world is deteriorating and the threat landscape in which we have to conduct our operations is becoming more complex and challenging. Among other things, the complexity is increasing partly because many independent factors influence societies and conflicts. War far from home can no longer be separated from security close to home and alongside small-scale expeditionary action ('wars of choice'), the large-scale conflict ('wars of necessity') is another scenario we should consider seriously. Our military capabilities evolve with these changes and the threat landscape. In 2019, the first F-35s arrived in the Netherlands, followed shortly after by the MRTT military transportation airplane. Later on, we will modernise our helicopter fleet and tactical air transport fleet. We expect unmanned systems capability somewhere in the following years.*
>
> *Effective use of these new capabilities requires new ways of thinking, working, directing, supporting and fighting. The Air Force must therefore change to become a data-rich, networked, challenging and hybrid environment to remain relevant: we call this the 5th Generation Air Force. With the decreasing cost of exponential technology, the availability of technology has become democratised. As a result, our Air Force is losing its technological edge. Therefore, people have to make the difference. This means that the 5th Generation Air Force is all about the professional, a highly educated and well-trained person who, as an individual, is capable of making a real difference. Especially in the 5th generation, the Air Force employee remains crucial. The 5GAF employee (military and civilian) must provide the competitive advantage* (Royal Netherlands Air Force, 2017a).

Regarding the Air Force employee, the leaflet goes on to say:

> *More than ever, the Air Force is about making a difference. We are not the only ones thinking about the future. Our adversaries have similar weapon systems, assets and ideas. People must make the difference, so employees are crucial to the 5GAF. [We need] well-educated, well-trained team players who can make a real difference, who want to get the job done and make the right decisions every day, individuals who want to change and keep on improving themselves* (Royal Netherlands Air Force, 2017b).

With regard to the management of the organisation:

> *We run our organisation on the basis of trust and accountability. Responsibility and power are placed as low and integrally as possible in the organisation. We set clear frameworks and assignments and give more freedom and space to our employees. There is more room for self-organisation, horizontal connection and multidisciplinary cooperation rather than top-down control. With the shift of focus onto effectiveness, the emphasis will be on organising differently, developing other desired behaviours and recalibrating our values and norms. We need curious and respectful people, who are driven and understand that at their own level they are contributing to delivering airpower. Managers must be able to build trust quickly, set clear goals and coach their people. We are striving for a culture of ownership, inspiration, external focus, continuous learning and connecting. We will harness our employees' talents better so that we support them in doing what they are good at instead of restricting or limiting them. This is how we can ensure that the RNLAF is willing and able to take on more responsibility*(Royal Netherlands Air Force, 2017a).

As this information on the 5GAF shows, in order to stay relevant and gain the competitive advantage it needs, the RNLAF not only wants new weaponry but also a different management style that will foster different behaviour in employees. Specifically, a 5GAF employee is a professional aiming for [self-]improvement and innovation, who uses the freedom given to make the right decisions to get the job done.

3. Theoretical background

In 19th century safety science, two themes dominated the field: the need to study the cause of accidents and the need to create and maintain safety rules and practices (Dekker, 2019). Since then, the concepts of behaviour, performance and rules have been linked to gain control over safety, with the focus on the topics of safety culture, just culture and compliance.

In the domain of safety science, two widely acknowledged and distinct safety paradigms can be distinguished (Dekker, 2019) (Hollnagel et al., 2015) which interpret these concepts and topics differently.

The first paradigm, often referred to as Safety I, is seen in the timeframe 1900-1999. Influenced by great names such as Frederick Taylor (1911) and James Reason (1997), the first paradigm considers safety as the absence of negative events, such as incidents and accidents (Dekker, 2019) (Woods et al., 2010). Seeing safety like this leads to defining the cause of negative events to either eliminate or contain them. Among the causes are humans behaving badly or not following the rules, and the underlying safety culture (Hollnagel et al., 2015). In Safety I, preventing negative events means intervening on the level of individual behaviour and/or at the level of the organisation. In either case, it is where things go wrong that needs to be rectified (Hollnagel et al., 2015). Typical intervention strategies are: safety management systems, compliance management, training; all strategies aimed at standardizing human behaviour within a system (Dekker, 2014b). However, as Hollnagel et al say: "*Trying to achieve safety by constraining performance variability* (e.g. standardization) *will inevitably affect the ability to achieve desired outcomes as well and therefore be counterproductive"* (Hollnagel et al., 2015).

Looking at the accident track record of a high-risk industry, organisation or company, it is evident that the number of accidents has seriously decreased since 1900, due to all the developments in training, rules, regulations and procedures, the introduction of safety management systems, ongoing development of the design and the interface with the human operator, and so forth (Dekker & Pitzer, 2016). Also evident is that after a certain point – about 10^{-6} or 10^{-7}, indicating that one in a million to one in ten million exposures results in a deadly occurrence (Amalberti, 2001) – any more of these interventions would not result in any further decline in the number of accidents. (Dekker, 2019).

This revelation lead to a new safety paradigm, often referred to as Safety II (Hollnagel, 2014) or Safety Differently (Dekker, 2014b). This paradigm recognises that because of technological innovation and increasing performance demands, humans work in increasingly complex systems which they no longer can fully understand or control, leading to uncertainty (Hollnagel, 2014). The fact that so few accidents happen is seen as a result of performance variability, meaning the ability of humans to respond and adapt to uncertainties (Hollnagel et al., 2015). Safety is therefore seen as the presence of a person's

positive capacity to adapt to variable conditions of the system in which they work, not the absence of negative outcomes (Dekker, 2014a) (Dekker & Tooma, 2021). In other words, the organisational system is inherently unsafe and employees create safety by adapting to the complexity of the system (Dekker, 2019) (Hollnagel, 2014) (Hollnagel, 2018). Negative human behaviour is driven by system failure (Dekker, 2002b).

Instead of just finding the causes of negative events, this Safety II focuses on understanding the dynamic, complex nature of a system and the adaptations employees make within that system. This obtains insight into the results of an organisation, whether positive or negative. While acknowledging the safety results accomplished by all the measures of the first paradigm, now the focus has shifted from preventing negative results to enhancing adaptive behaviour in operators to ensure as many positive results as possible (Hollnagel et al., 2015).

As previous research shows (Hale & Borys, 2013), both paradigms are needed to ensure safety in aviation. This is why the Safety II / Safety Differently paradigm does not discard Safety I standardisation, rules and compliance (Hollnagel, 2014). However, despite the positive impact of Safety I measures, Safety II rejects the way Safety I views operators and their behaviour as liabilities and hazards (Hollnagel et al., 2015). Instead, Safety II sees them as professional experts who need discretionary space for adaptive behaviour to achieve an expected safety level (Dekker, 2014b).

Three key concepts of both paradigms are relevant to my research: safety culture, just culture and compliance vs. adaptation, which I will go on to describe briefly. These three concepts take up a significant part of the literature focused on enhancing safety and therefore, put together, they might provide a solid explanation for not only the safety an organisation achieves but also the stalemate and plateauing results an organisation meets when trying to further enhance safety (Dekker & Pitzer, 2016).

3.1 Safety culture

The concepts of organisational culture and safety culture are well studied, discussed and debated (Aburumman et al., 2019) (Cox & Flin, 1998) (Falconer, 2006) (Ferraro, 2000) (Filho et al., 2010) (Glendon & Stanton, 2000) (Glesner et al., 2020) (Grote, 2007) (Guldenmund, 2000) (Guldenmund, 2010) (Health & Safety Executive, 2005) (Helmreich, 1998) (Johnsen et al., 2003) (Krause & Henshaw, 2005) (Le Coze, 2019) (Lund & Aarø, 2004) (National Aeronautics and Space Administration, 2003) (Neal et al., 2000) (Obadia et al., 2007) (Parker et al., 2006) (Persson, 2012) (Schein, 2004) (Sexton & Klinect, 2017) (Soeters & Boer, 2000) (Sorensen, 2002) (Wiegmann et al., 2002) (Yeoman, 2011) but to say a consensus has been reached is far from academic reality. Safety culture is usually seen as a specific, fundamental part of organisational culture (Hayward, 1990) (Iordache et al., 2016). It is commonly agreed the interest in safety culture dates from the Chernobyl accident in 1986 (Health & Safety Executive, 2005) (Obadia et al., 2007) (Patankar et al., 2012) (Thaden & Mitchell-Gibbons, 2008). In the aftermath of this accident, the International Nuclear Safety Advisory Group (INSAG, 1991) introduced the term ‘safety culture’. Since then, it has been seen as the key to improve safety (Antonsen et al., 2017). Safety culture is a vast growing research field, yet there is no consensus amongst researchers as the exact definition of the concept and its precepts (Choudhry et al., 2007) (Edwards et al., 2013) (Guldenmund, 2000).

Nonetheless, every definition of safety culture focuses on the priority given to safety, the position it holds within an organisation and the impact the culture has on safety (Cooper, 2000) (Cox & Cox, 1991) (Lee, 1998) (Merritt & Helmreich, 1996) (Wiegmann et al., 2002).

There are various views on the relationship between safety culture and safety. Some researchers find that it has an impact on the safety of an organisation in that incidents and accidents occur when the organisation lacks a good safety culture (Rubin et al., 2020) (Vuuren, 2000) (Wang, 2018)(Rubin et al., 2020) and these events (Stolzer et al., 2011) (Zwetsloot et al., 2017) or their impact are reduced when it has a good safety culture (Berglund, 2020). An organisation's safety culture also seems to influence the extent to which safety procedures are followed (Antonsen, 2009). It appears that culture can have a negative impact on safety attitudes and behaviour and shared safety values can predict at-risk behaviour (Watson et al., 2005). For instance, Cavazza and Serpe (2009) found that a positive safety climate results in workers being more inclined to use personal protective equipment and show less unsafe behaviour.

Other researchers find safety culture has a limited predictive ability (Antonsen, 2009) (Dekker, 2019). Error is not caused by incompetent employees but is the result of a system not adapting to the changing world around it (Dekker, 2007a) and is a by-product of the pursuit of success (Dekker, 2006). A safety culture, therefore, is not without error. It recognises that besides making mistakes, people create safety through practice (Dekker, 2007b)(Dekker, 2007c). Counting errors often leads to the wrong conclusion. When an organisation does not count errors because none have been made, this does not mean that the organisation is safe and vice versa (Dekker, 2009).

Studying safety thus need a shared frame of reference which includes topics such as politics, power and conflict. This can shed light on, for instance, the different perspectives on safety between management and work floor (Antonsen, 2009a). Enhancing safety through the influence of safety culture seems unlikely, because safety culture is created locally in a daily interaction which cannot be controlled (Antonsen, 2009c). The organisation should not try to change its safety culture but try to understand it in order to change the organisation to enhance safety (Antonsen, 2009c). In other words, safety culture can be a valuable concept, but the practical value is considered debatable (Le Coze, 2019).

3.2 Just culture

Related to the concept of safety culture is the notion of just culture. As espoused in the safety domain, learning from accidents, incidents or near misses allows organisations to improve safety (Dekker, 2006) (Reason, 1997) (Woods et al., 2010). Learning is particularly key to organisations operating in high-risk environments (Catino & Patriotta, 2013). Vital to this learning is safety-related information (Dekker, 2002a), that operators involved in occurrences possess, as they are the closest witnesses when these negative events unfolded. How an organisation responds to an occurrence determines the quantity and quality of information that runs freely throughout the organization. It determines how safe people feel to share their concerns.

The safety I paradigm led to two visions of safety enhancement – behaviourism and reliability engineering – which influenced the way information is shared within an organisation. On the one hand behaviourism focuses on individual behaviour, seeing it as a liability and compliance as a means to

contain this hazard. Operators are held accountable for non-compliance, and individuals are blamed for negative events resulting from the logical consequence of noncompliance. This makes sharing of information difficult because operators can be punished according to what they could reveal.

On the other hand, reliability engineering focuses on the liabilities in a system and engineering as a means to make a system as reliable as possible. Since this system is unreliable until all the limitations have been engineered away, operator behaviour is not a cause but a symptom of underlying systemic problems (Reason, 1997). To ensure that safety problems are reported, the organisation should provide safety information and, vitally, must hold a no-blame attitude, except in cases of gross negligence (Reason, 1997). These two outcomes, blame and (mostly) no-blame were considered irreconcilable.

Reconciling the two entails a compromise between a fear of blame-free approaches and need for accountability and the willingness of employees to voluntarily report safety-related issues without being punished for it (Dekker & Breakey, 2016) (International Civil Aviation Organization, 2018), Reason (1997) introduced the concept of just culture to solve this tension between blame and no-blame. Just culture is one of the four sub-cultures of safety culture. It is seen as an atmosphere of trust, where events are shared and reported (Kováčová et al., 2019) and in which there is clarity on what is and is not acceptable (Balcerzak, 2017) (EUROCONTROL, 2006) (Reason, 1997). Such sectors as aviation, healthcare, nuclear organisations and the railway have embraced the concept of just culture (Heraghty et al., 2021).

Using Reason's (1997) supportive decision tree as an analytical tool, shown behaviour is scaled either acceptable or unacceptable depending on how the individual answers the questions. An individual can be both blamed and exonerated for whatever negative event happened. Either way, learning occurs, be it from understanding how operators get caught up in an undesired event (no-blame) or by making clear what constitutes unacceptable behaviour (blame).

However, this concept of just culture has not escaped its critics. Regardless of the outcome, blame or no-blame, the decision tree judges individual behaviour, leading one to question who does the judging and who decides what is acceptable or not (Dekker, 2009). Just culture models designed to determine whether an event is classified as a mistake or as negligence and that draw a clear line between acceptable and non-acceptable behaviour (Marx, 1997) (Reason, 1997) lead to nothing less than a retributive just culture (Dekker, 2016).

The retributive approach is a fairly quick, clear-cut process. The consequences of errors are clear to everyone (International Civil Aviation Organization, 2018). Acceptable and non-acceptable behaviour are clearly defined (Reason, 1997) and balance is restored by holding someone accountable (Dekker, 2016). As long as the consequences are proportionate to the actions, retribution is seen as just (Dekker, 2016). However, if the consequences are perceived as unjust this can be counterproductive to the intended open culture (Brborović et al., 2019) (Lawrenson & Braithwaite, 2018) (Lipira & Gallagher, 2014) (Reason, 1997).

Acceptable and unacceptable behaviour are not fixed categories but are determined by time, context and the person making the judgement (Alicke et al., 2008) (Dekker, 2003b) (Dekker, 2009). When a judgement seems unjust, the information flow is hampered because people might try to hide their mistakes to avoid negative consequences (Dekker, 2007e). The fear of repercussions (retribution) can lead to quick confessions with minimal disclosure motivated by the thought that the sooner the act is reported, the more often it is considered brave the consequences will not be that bad. The emphasis is

on avoiding consequences instead of sharing information (Dekker & Laursen, 2007). Furthermore, punishing and learning are fundamentally at odds. Whereas punishing has to do with continually believing the system is safe, learning has to do with changing these beliefs (Dekker, 2003a). Punishing means moving away from the adverse event instead of confronting it (Dekker, 2007a), providing the illusion of control and fear instead of awareness (Dekker, 2007b).

As opposed to retributive just culture, restorative just culture counters the negative effects of retribution. The focus of a restorative just culture is on repairing (restoring) and learning (Dekker, 2016) (Dekker & Breakey, 2016) (Goodstein & Aquino, 2010), on gaining knowledge of the context and system and finding the answer to the question of how to move forward (Parkinson & Roche, 2004). People are still held accountable, but what needs to be done is based on a broadly-based decision involving everyone affected by an event, delivering more satisfying results for all concerned (Braithwaite, 2002).

Restorative just culture appears to have a positive impact on safety in organisations. Turner et al. (2020) describe how implementing a restorative just culture increases trust, shared accountability and learning. The focus on repairing and learning delivers richer information (in both quantitatively and qualitatively) to detect near misses, system solutions and a safer workplace (Dekker & Breakey, 2016; Bell et al., 2011). Kaur et al. (2019) find that introducing a restorative culture seems to lead to fewer suspensions and disciplinary measures, reduced absences due to illness, and increased disclosure and cost effectiveness.

3.3 Compliance vs. adaptation

My third topic of interest is compliance. The value and desirability of rules intended to enhance safety (Gherardi, 2018), specifically in dangerous situations (Hale, 1990), are well known (Weichbrodt, 2015). Rules are a necessary resource in organisations. They help people remember the steps in a task under challenging circumstances. They are used to help educate and train people for their jobs, ensuring people can cooperate effectively. They are used for design and planning purposes and as a means to identify variances in behaviour that are deemed unacceptable (Hale & Borys, 2013). This last aspect means that rules provide the means to appoint accountability in case something goes wrong (Almklov et al., 2014) (Dekker, 2014c) (Mascini, 2005).

The literature distinguishes three types of rules (Hale & Swuste, 1998):

- rules that define the goals that must be achieved (performance goals);
- rules that define the process to be followed or decisions to be made in order to get the job done (process rules;
- rules that define what specific behaviour is expected of employees (action rules) (Hale & Borys, 2013).

The discretionary space employees have lessens from performance goals to process rules to action rules (Hale & Swuste, 1998).

Implementing rules and procedures have paid off going by the reduced number of deadly accidents (Amalberti, 2001). However, the 20th century saw an ever-increasing number of rules and procedures, and this number is still rising (Dekker, 2014c), for action rules and with ever-diminishing autonomy and discretionary space for workers (Dekker, 2020).

Because rules and procedures provide standard means of operating across different situations, the assumption within those rules is that the work itself is relatively static and predictable (Biermann-Teuscher et al., 2022). This view fits well with action rules. But rules and procedures are different from work-as-done (Hollnagel, 2012). The working environment is often complex, dynamic and non-linear whereas rules and procedures work well in routine and predictable situations (O'Keeffe et al., 2015). Because it is hard to capture the dynamics of daily work in rules and procedures (Biermann-Teuscher et al., 2022), this has led to the assumption that there is only one best practice, as espoused by the rules (O'Keeffe et al., 2015). But when the situation changes, the rules may no longer provide the best solution (Biermann-Teuscher et al., 2022).

Research shows that rules and procedures sometimes contradict one another(Bye & Aalberg, 2020), do not match the work to be done (Blazsin & Guldenmund, 2015) (Carvalho et al., 2018), can be too complex for use (Antonsen et al., 2008) and actually prevent workers from doing the work (Bieder & Bourrier, 2013) (Biermann-Teuscher et al., 2022) (Damoiseaux-Volman et al., 2021) (Halbesleben et al., 2008). As a result, rules and procedures are ignored on many occasions (Gray, 2006) (Lawton, 1998) (Tucker, 2004) (Tucker et al., 2020) and because of this ill-fit. they are not enough to create safety (Dekker, 2003a). Moreover, safety requires adaptive behaviour because work takes place in a context of limited resources, multiple goals and pressures as well as surprises, constraints and novelty which rules and procedures cannot fully describe (Dekker, 2003a). In fact, because it is impossible to anticipate every possible work situation, operators find it necessary to break the rules and depend on informal work practices (Bye & Aalberg, 2020) (Nordlöf et al., 2015) to accommodate the variety in work situations (Woods & Shattuck, 2000).

Common sense is found to be important to get the work done (Nordlöf et al., 2015), sometimes even more important than complying with safety rules (Thorvaldsen, 2013). In other instances it might actually mean breaking rules and procedures (Desai, 2010). Given these problems with rules and procedures, there is wide acknowledgement of the difference between paper and practice (Damoiseaux-Volman et al., 2021) (Dekker, 2006).

Today, many organisations face an ongoing dilemma caused by the tension between two necessities, desired compliance and necessary adaptation (Hollnagel, 2014) (Mendoza et al., 2020). To get the job done, organisations need both rules and procedures as well as the common sense and informal work strategies employed by operators (Gherardi & Nicolini, 2002). This dilemma (Biermann-Teuscher et al., 2022) can be represented by Dekker's two distinct models (2014a). Model 1 fits the Safety I paradigm and is based on the idea that following procedures means being safe (Dekker, 2014b) and accidents are considered proof that rules and procedures should be followed (Mascini, 2005). Any gap between rules and reality needs to be resolved by matching reality to the rules (Hale & Borys, 2013).

Model 2, resembling the Safety II paradigm, does not consider that procedures guarantee safety. Rules cannot describe all the possible circumstances (Hale & Borys, 2013). Violating or adapting procedures is necessary because multiple, conflicting goals must be achieved (Dekker, 2014b). Employees get the work done by being more efficient or more thorough than expected, capable to dealing with a lack of resources (Ashour et al., 2021), able to innovate (Dekker, 2018)] or use work arounds (Tucker et al., 2020). Any gap between rules and reality needs to be resolved by changing the rules so that they match reality (Hale & Borys, 2013).

These two paradigms have a distinct impact on the framing of rules and violations. Model 1 thinking considers rules cast in stone and violations the first step in creating an accident. Model 2 thinking considers rules as guidelines to be adapted and complemented with expert knowledge when the situation demands (Hale & Borys, 2013).

3.4 Relationships between safety culture, just culture and compliance

The three topics (safety culture, just culture, compliance), although often studied separately, are interrelated. Rules and procedures are regarded essential elements of a safety culture (Cooper, 2000) (Guldenmund, 2000). The response to violating rules and procedures shows the just culture of an organisation. Table 1 summarises the way the three topics relate to the two safety paradigms.

Table 1: Summary of the two paradigms and three topics in this research

Safety I:	**Safety II/Safety differently:**
Safety culture: the shared notion of safety and, when measured and understood, a means to explain incidents and predict and enhance safety.	Safety culture: a shared frame of reference but created locally. A valuable concept to be understood so the organisation can be changed to match the local circumstances.
Just culture: focuses on people, whether or not blaming is appropriate, and uses consequences to restore balance. This form is called retributive just culture.	Just culture: focuses on restoration, what is needed to repair the damage and move forward. This form is called restorative just culture.
Compliance: following the rules leads to safety. Any gap between paper and practice needs to be solved fast by changing practice. This notion is called model 1 or the classical paradigm.	Compliance: as rules can never be complete or perfect, safety results from employees adapting to the situation. This notion is called model 2 or the constructivist paradigm.

The three topics become increasingly visible in research into undesired outcomes, like incidents and accidents. For instance, research shows that an accident often results in politicians, press and public demanding to know the cause and taking appropriate measures (Mascini, 2005). Because organisations are eager to show that they take accidents seriously and want to restore lost trust as soon as possible, accidents are often blamed on rule breaking with punishing the rule breakers as a logical consequence (Mascini, 2005). Studies have found that operators who broke a rule followed by an accident were punished more severely than operators who broke a rule that was not followed by an accident (Mascini, 2005). These result show the relationship between just culture and compliance; compliance is used as both a promise to the outside world that future occurrences will be prevented and as a measure to determine the amount of blame appointed.

Regarding compliance, a judgement like this says operators could know beforehand that an accident would happen because they did not follow the rules (Hale & Borys, 2013), confirming the model 1 vision that rules are good, violations are bad (Mascini, 2005), and safety can be controlled beforehand (Biermann-Teuscher et al., 2022). Because events are considered controllable, blaming individuals for the outcome seems to make sense (Lupton, 2006), again showing the relationship between just culture and compliance thinking.

From a model 1 perspective, more rules and increased pressure on compliance seems a good way to prevent future accidents. Research shows that many organisations come up with extra rules after an accident (Mascini, 2005). With this focus on compliance, as in model 1 thinking, operators become increasingly responsible for their own safety (Biermann-Teuscher et al., 2022) (Morris et al., 1999) and the target of new rules and procedures (Gray, 2006). Safety culture can explain why rules and procedures are adhered to or not (Antonsen, 2009b), indicating the relationship between compliance and safety culture research. A focus on accountability can result in people hiding occurrences instead of talking about and learning from them (Biermann-Teuscher et al., 2022), indicating, as Reason (1997) has pointed out, the relationship between just culture and safety culture.

Another situation that shows the relationship between the three topics is the employees' need to prioritise. Operators are regularly confronted with a necessary trade-off between productivity and safety (Hollnagel, 2009) (Nordlöf et al., 2015), a trade-off that is made based on their knowledge and expertise. It has been found that workers prioritise production over safety procedures to make sure the work gets done (Antonsen, 2009) (Atak & Kingma, 2011) (Brown et al., 2000) (Nordlöf et al., 2015) (Walker, 2008). They adapt existing rules on a daily basis, applying 'safe-enough' strategies (Bye & Aalberg, 2020) (Nordlöf et al., 2015) (O'Keeffe et al., 2015). Research shows that in many instances the workarounds executed by operators to get the work done are safe, sometimes even safer than following the rules (Mascini, 2005), indicative of a relationship between compliance thinking and safety culture research. Workers adapt to the local circumstances and apply their own informal work practices alongside rules and procedures (Biermann-Teuscher et al., 2022). Increasing safety can be considered something that emerges from operators' adaptations to local circumstances as opposed to solely regulated by rules and procedures (Halbesleben et al., 2010). In this sense it is based on the understanding of the situation by professional employees, alongside their knowledge of rules and procedures (Weichbrodt, 2015). Not recognising the operators' contribution, based on their knowledge and experience, is unjust treatment (Biermann-Teuscher et al., 2022). Moreover, due to the common retributive response towards violations, these workarounds remain rather hidden (Hale & Borys, 2013), which means precludes the possibility of learning and improving from them (Biermann-Teuscher et al., 2022) and the organisation stays unaware of exactly how the work gets done. Again, these results are indicative of the relationship between the safety culture, just culture and compliance thinking.

4. Rationale fort his dissertation

The three themes of safety culture, just culture and compliance provide a common thread in the safety documents of the MoD appearing in the past six years.

4.1 Safety culture within MoD

The notion of a 'good safety culture' in the MoD is quite apparent. For instance, in 2017 the Dutch Safety Board wrote in an investigation report:

> *"Safety culture and safety awareness are important pillars for a safety defence organisation […]. Invest in [an] organisational culture in which management is receptive to criticism from employees. […] Encourage communication about safety risks to provide a broad safety awareness within the defence organisation."* (The Hague: Onderzoeksraad voor de Veiligheid, 2017)

The MoD responded in a letter to the Dutch House of Representatives:

> *"The Dutch Safety Board is right to state that safety culture and safety awareness are important aspects of a safe defence organisation."* (Ministerie van Defensie, 2017)
>
> An external committee was installed to investigate how the shortcomings, identified by the Dutch Safety Board, arose in the MoD. In 2018 this committee concluded in the final report:
>
> *"[Recent] budget cuts have made an appeal on 'can do' mentality […] which has resulted in a gradual but steadily growing culture that accepts more risks, contrary to rules or procedures […]. The safety culture […] needs close attention to ensure structural safety in daily operations."* (Commissie Van der Veer, 2018)

To which the MoD responded to the House of Representatives:

> *"We fully agree with the conclusions [of the committee] and will comply with the recommendations. Enhancing safety management explicitly requires […] a change in behaviour and culture. […] Safety is not a sufficient part of the […] organisational culture to ensure safe task execution."* (Minister en Staatssecretaris van Defensie, 2018)

This promise was reflected in the safety plan of Central Staff that included 40 measures to enhance MoD safety:

> *"We will achieve this by investing enormously in the organisational ability to learn, the safety culture, […]. The MoD has a unique organisational culture [which has] proven its worth again and again during deployments. [However, it] cannot result in accidents or […] undesired behaviour. […] We will strive for a culture [that] actively encourages speaking up but […] gives a firm response when necessary. A culture in which we learn from our mistakes. Awareness of the undesired consequences of behaviour is the main focus. It is up to all of us to ensure an organisational culture that makes safety an integral part of what we do and think."* (Ministerie van Defensie, 2018)

The specific measures to achieve the desired organisational safety culture are:

- Conduct behavioural interventions based on risk analyses or research;
- Write and implement a safety communication plan;
- Hold annual symposia and workshops on psychological and physical safety;
- Commanding officers/management devote daily attention to exemplar behaviour and speaking up;
- Add 'Willingness to report' to the existing Code of Conduct;
- Enhance knowledge of safety with an information campaign;
- Conduct initiation rituals according to rules;

- Enhance attention for moral fitness in education, training and daily operations;
- Continue the 'Contradiction, yes please!' campaign (Dutch: 'Tegenspraak, ja graag!');
- Mount campaign to enhance the willingness to report occurrences.

To ensure that the MoD lived up to the promises made in the safety plan, a committee was assigned in 2018 to monitor the implementation of the measures. In 2021, their final report stated:

> *"The safety plan pays explicit attention to behaviour, cultural change and the role of management. In the first report (2019) the committee concluded that the safety culture amongst operators is better than expected. That conclusion has not changed since."* (Visitatiecommissie Defensie en Veiligheid, 2021)

As these excerpts show, there was outside pressure to enhance the safety culture in the defence organisation (Dutch Safety Board, Van der Veer committee), even if the concept remained vague. In response, the MoD committed itself to clarify the vague concept to enhance safety. Remarkably, the committee installed to monitor the implementation of the safety plan was far more positive about this culture when it came to the work floor than the other committees and MoD were.

4.2 Just culture within MoD

On the notion of just culture, the following excerpts come from the reports of the committee installed to monitor the safety plan implementation. In 2019 they stated:

> *"It is therefore of utmost importance that these mistakes can be made in a psychologically safe environment. Mistakes can be discussed, with learning taking priority. Reporting (near) incidents supports a learning organisation. Whenever blaming and learning questions are confused, the willingness to report an occurrence will drop. This has a negative impact on a learning organisation. The committee sees that the MoD is currently struggling to find a balance between investigating to learn and pressure from outside the organisation to appoint blame for occurrences."* (Visitatiecommissie Defensie en Veiligheid, 2019)

The next year, the interim report stated:

> *"Research shows that over ten percent of MoD employees report a psychologically unsafe work environment. On the other hand, the committee has observed a great deal of trust amongst operators, which is an important indicator of a psychologically safe work environment. The MoD desires an organisational culture in which employees can be themselves without fear of exclusion and which accepts reporting of misconduct."* (Visitatiecommissie Defensie en Veiligheid, 2020)

One year later, the final report stated:

> *"The MoD desires to be a learning organisation. Too often the committee has seen cases in which mix up learning and punishment. In the case of incidents, ensure an investigation into the facts, not the mistakes. Criminal law should be a final resource, only to be used once it is established that other measures do not lead to the desired effect. [The] questions of learning and blaming are mixed up. Learning can only happen if the 'why' question is answered. A one-*

dimensional answer to cause-result reasoning is not enough. […] There needs to be room to conclude that no reasonable measures could have prevented the occurrence.

"[…] In earlier reports the committee concluded that the MoD is struggling to find a balance between investigating to learn and pressure from outside the organisation to appoint blame for occurrences. […] The committee also concluded that in an organisation focused on learning, employees should be able to make mistakes. Expectations from outside the organisation to find the underlying cause of an occurrence do not support this aspiration. The committee senses a hesitation in the MoD to communicate about occurrences. When it comes to publicity about awkward occurrences the MoD recognises this defensive attitude and notes a dilemma. The MoD wants to learn from mistakes, but at the same time feels compelled, possibly by the outside world, to also ask the 'blame' question." (Visitatiecommissie Defensie en Veiligheid, 2021)

These excerpts show an underdeveloped restorative just culture. Employees do not feel psychologically safe. Safety investigations seem to contradict the aspiration to learn and blaming is a likely outcome. This is also apparent in the story at the beginning of this introductory chapter.

4.3 Compliance within MoD

As for compliance, the following statements are found in MoD documents:

"The desired end state is a culture in which the focus of attention is not only on employees following the rules (compliance), but also on employees acting proactively (intervention)." (Ministerie van Defensie, 2018)

"A proper way of organising safety is by implementing procedures that actually enhance safety. […] Imposing procedures that differ from daily practices constitutes a false sense of safety. Research shows that employees cannot comply with procedures that deviate from their daily routines." (Visitatiecommissie Defensie en Veiligheid, 2019)

"The committee acknowledges that the MoD is very skilled in making plans. It puts a lot of effort into planning, even when there is only a small risk. These plans are not connected and sometimes are so vast that questioning the balance between policy and practice is justified. According to the committee, safety plans can be limited in scope but also enriched with people, means and time. In other words, plans that can realistically be implemented. […] When implementing safety measures and communicating on risks, include insights into the risk-rule reflex. The MoD tends to implement vast, heavy measures (more rules, more supervision) to control a new risk or prevent an occurrence from happening again, often at high policy costs and diminishing improvement revenue, unclarity about responsibilities and unrealistic expectations. For example, the committee found the measure to ban holsters unsafe. When such a decision is made, a different risk appears: military personnel are to bring their guns without a holster. This begs the question of whether the measure to contain a risk (banning unsafe holsters) leads to a greater risk (unsafe transportation of guns). The point is that risks should be identified, weighed against other risks and possibly accepted based on

rational grounds. When deciding on measures, the pros and cons must be balanced. Make sure that the organisation has a balanced safety policy and that management takes on the responsibility to accept certain (smaller) risks. Make sure there is a difference between operational safety and safety in business operations. […] Internal transport [of dangerous goods] within the MoD does not fall under the law regarding the transportation of dangerous goods. Based on a risk assessment the MoD is allowed to set its own priorities. But the cases provided to the committee show that this discretionary space is not used to consider doing something safely for the sake of daily life, business operations or military exercises." (Visitatiecommissie Defensie en Veiligheid, 2021)

These excerpts show that the MoD, although seemingly unaware of the contradictory paradigms, recognises that it needs to think in a combination of model 1 (compliance) and model 2 (proactive intervention = adaptation). The committee's opinion is that the MoD stands to benefit from rules and procedures that match daily practice and produce less paperwork, which is more of a model 2 focus.

4.4 Rationale for the research

The MoD documents and the scientific literature on safety culture, just culture and compliance provide the rationale for the research described in this dissertation.

The Dutch MoD struggles with the concept of safety culture. It promises to enhance it, but six out of ten measures in the safety plan focus on providing information to employees. The remaining four measures focus on compliance or changing individual behaviour. In all cases the relationship with safety is unclear, as it also is in the scientific literature. The notion of safety culture is considered important because the MoD/RNLAF wants to prevent accidents. Furthermore, the MoD wants to have a culture of openness, with employees speaking up and learning, yet is unable to move beyond accountability and blame. It desires both compliance and proactive intervention but does not know how to achieve this.

What seems clear from the committee documents is that operators create safety in the MoD, since the safety culture, whatever it may be, is better among operators than might be expected. Among equal colleagues there is horizontal trust. Trust is only problematic in the vertical, hierarchical organisation. It is contradictory that MoD expects operators to intervene proactively while, at the same time, these operators feel the lack of psychological safety and get blamed for undesired outcomes. Fear of repercussion most likely withholds operators from adapting unless they can hide their informal work practices.

Research on the border of the three domains (safety culture, just culture, compliance) and how they are related provides scientific insight into the underlying patterns. Besides adding to the theoretical knowledge in these domains, this research supports a struggling MoD, and therefore the RNLAF, in its desire to enhance safety. Given that the transition to a Fifth Generation Air Force requires both management and employees to behave differently in dealing with increasing complexity, these insights will also support the RNLAF in its ambition to become a 5GAF.

5. Research question and sub-questions

The RNLAF slogan is 'mission first, safety always'. However, documents show that the MoD/RNLAF is struggling with the concept of safety, which makes one wonder how 'safety always' is achieved. The aim of this dissertation is to understand how the RNLAF can achieve both safety and effectiveness in order to support not only the operators performing in complex and dynamic situations but also the staff in their ambition to become a 5GAF.

The following question encapsulates the focus of this dissertation:

How can the safety and effectiveness of the RNLAF be described and enhanced?

Five sub-questions, all focused on understanding safety, examine various aspects intended to enhance safety:

1. What is the predictive value of a safety culture assessment regarding the accident-proneness of an operational squadron? (Chapter 2)

The MoD, and therefore the RNLAF, is committed to enhancing the safety culture in order to prevent negative outcomes. Although the literature is inconsistent on the notion of safety culture and its predictive value, I conducted the first study to see if adding more detail to a safety culture assessment would enhance its predictability and therefore its value as a method. If safety can be predicted using a safety culture assessment, then the RNLAF knows what to focus on when trying to prevent occurrences, and will not be wasting time, effort and resources on the wrong matters. Both preventing occurrences and appropriately distributing resources enhances the effectiveness of the organisation.

2. What are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013)? (Chapter 3)

Because the first study focuses on lack of safety – the predictive value of a safety culture assessment regarding the accident-proneness of a squadron – this second study focuses on the presence of safety. Since the realisation of safety is influenced by the way rules are viewed, that is, model 1 and model 2 thinking, this study focuses on the impact of these two models on the working environment and safety of operators in an operational squadron. Discovering which paradigm regarding rules and compliance best suits the RNLAF can ensure that operations and task execution go as smoothly as possible, which will contribute to the effectiveness of the organisation.

3. Why do differences between the way work is done and prescribed endure despite their undesirability? (Chapter 4)

As a result of the tension between the desire for compliance (model 1) and the need for adaptation (model 2) undesirable gaps between paper and practice are pretty common. Because these gaps hamper both safety and effectiveness, the question this study seeks to answer is why these gaps remain. If the reasons for enduring gaps are clear, solving them becomes possible, enhancing both safety and effectiveness.

4. What complexities arise when attempting to foster a restorative just culture in a mostly retributive organisation? (Chapter 5)

The solutions operators apply when dealing with conflicting goals (resulting in gaps between paper and practice), remain hidden from the organisation outside the squadron. Just culture determines the amount of transparency and communication within an organisation. A restorative just culture fosters more openness than a retributive just culture. Because the RNLAF is mostly retributive, this study looks at how a restorative just culture can be fostered within a retributive culture. A restorative just culture fosters more transparency, which means that it is easier to discuss local ingenuity, further enhancing the safety and effectiveness of the organisation.

5. Are micro-experiments enough to reduce the enduring gaps between work-as-imagined and work-as-done? (Chapter 6)

The gaps between paper and practice that arise when operators apply local ingenuity to deal with conflicting goals are considered undesirable by the organisation. This study examines how the gap between the way work is done and the way it is prescribed can be reduced practically by combining micro-experiments with Hale and Borys' (2013) approach, which places monitoring and adaptation of rules central to its management process. Reducing the gaps between paper and practice eases tasks execution and therefore enhances both safety and effectiveness of the organisation.

6. Outline of dissertation

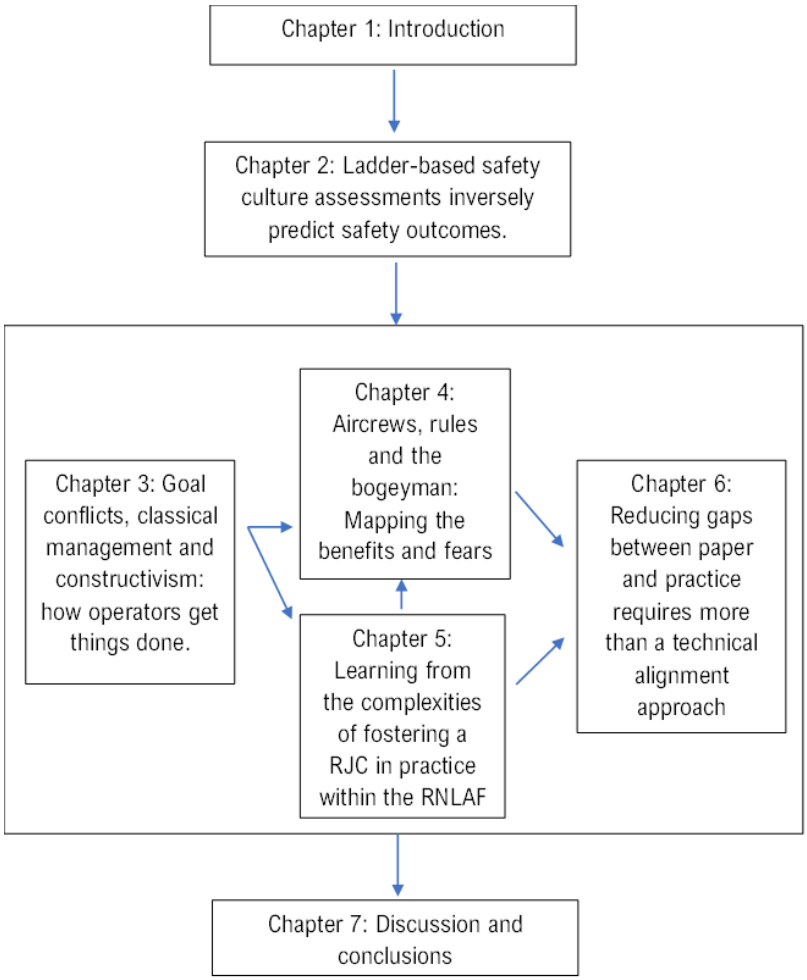

Figure 1: Outline of dissertation

Besides this introductory chapter, this dissertation contains another six chapters.

Chapter 2 presents a study that seeks to extend previous research by Antonsen (2009) that showed an inverse relationship between a quantitative safety culture assessment (SCA) and incident investigation findings. My study draws a similar comparison but contains more detailed information in the SCA. For data collection, instead of a questionnaire with a 5-point Likert scale, my study used a safety culture ladder method containing 28 items. Each item encompassing five detailed work descriptions representing the five safety culture maturity levels. Respondents from an operational squadron chose the description best fitting their work situation. By sheer coincidence, this same squadron experienced

an accident just a few weeks after the assessment. This provided a rare opportunity to compare the result of the accident investigation with the SCA. As with Antonsen's research (2009), my study also found that the investigation findings not only do not match the SCA results, indicating that the SCA has limited predictive value, but also shows an inverse relationship between the two sets of findings. If the desire is to gain insight into the state of safety within an operational unit, a ladder-based SCA does not seem to be the right choice.

Chapter 3 takes a first step to gain insight into the state of safety within a unit by digging deeper into the impact of rules and compliance on daily operations. Rule management on the one hand and the jobs needed to be done on the other hand leads to conflicting goals. When trying to realise goals, operators encounter two contradicting paradigms regarding rules and rule management. These are called model 1 and model 2 thinking (Hale & Borys, 2013). The RNLAF use model 1 thinking to the exclusion of model 2 thinking, as seen in the top-down approach of rules. Model 2 thinking is found in solutions applied by operators to solve goal conflicts. These solutions are dubbed local ingenuity, routines that are part of the regular repertoire of operators. These routines remain relatively invisible to management, are not originally intended and also not included in any existing rule base, even though they are part of the informal work system. However, local ingenuity is either limited to the caveats of existing rules or is not transparent and therefore suboptimal, resulting in overall suboptimal effectiveness. My study shows that noncompliance enhances goal realisation. A combination of model 1 and 2 thinking resulting in freedom-in-a-frame is needed to enhance both safety and effectiveness.

Chapter 4 expands on the research described in Chapter 3. Because in many cases it is impossible to follow the rules and get the job done at the same time, operators employ their own routines, dubbed local ingenuity. This results in work-as-done that differs from the work-as-prescribed. The focus on compliance in model 1 thinking on rule management and on rules that do not match the working situation seriously hamper task execution and the subsequent gaps between paper and practice are not resolved. I studied the stakes and interests that lie behind retaining these gaps, by mapping the fears and benefits associated with noncompliance. The research found a self-sustaining force field of interests, which simultaneously retain *and* deny noncompliance at almost all hierarchical levels. In addition, it uncovered routine invocation of a nondescript organisational force-to-be-feared in case noncompliance was discovered, reminiscent of a bogeyman. And although minor attempts are made to gain discretionary space and reduce the number of rules, these are unsuccessful, because the pressure to keep the promise of compliance makes change unlikely. Rather, the efforts should be directed at the bogeyman: the belief system that compliance equals safety, adhered to by politics and public and given lip service by military leaders.

Chapter 5 concerns a case study on the partly successful efforts to resolve the tension between achieving restorative justice in the predominantly retributive culture of the RNLAF. Restorative just culture seems to have a positive impact on safety culture in organisations. However, not much has been written on the complexities of developing this culture in an organisation accustomed to retributive just culture. Therefore, the objective was to identify and understand the complexities that arise when attempting to foster a restorative just culture in a mostly retributive organisation. The case study is in question was a 'safety standdown,' held in the RNLAF.

A safety standdown is a RNLAF intervention that cancels all flights and brings the aviation community of a squadron together to discuss the safety matters at stake. This particular safety standdown was

studied because of a spontaneous restorative response amidst an otherwise quite retributive setting. The findings show that four factors are of great importance for a restorative response: 1) a willingness to show vulnerability without being punished for this; 2) moral courage and a fair attitude, even if it means sacrificing one's own interests; 3) a management that does not punish commanding officers for responding restoratively; and 4) a desire to learn from the mistake instead of punishing for it.

Chapter 6 details a study that sought out to bring all the findings of the previous research together. This study used a combined approach of a rule management framework with micro-experiments to try to resolve the gap between the way work is done and the way work is described. It describes four micro-experiments of which three were completed. The focus was on adapting the existing rule base to the way work is done in practice. This study answers the question if micro-experiments are enough when trying to change rules. The findings show that micro-experimental processes are technically successful in showing the effectiveness of interventions and generating shop floor commitment to them. However, the findings also show that existing literature had not considered the difficulty of actually changing rules in order to close gaps between paper and practice. Finally the findings show that to be really successful, micro-experiments need to be supplemented by a focus on psychological safety, restorative just culture, psychological contracts and betrayal and moral disengagement.

Chapter 7 deals with the conclusions of the five described studies. The chapter brings together the insights of the earlier chapters and answers the central research question as well as the sub-questions. Reflecting upon the findings, I discuss how safety in the military and political context of the RNLAF can be enhanced. I indicate what is needed to become a 5GAF in which employees are the required competitive advantage. Finally I suggest how my insights add to the existing safety literature and what future research might entail.

References

Aburumman, M., Newnam, S., & Fildes, B. (2019). Evaluating the effectiveness of workplace interventions in improving safety culture : A systematic review. *Safety Science, 115*(February), 376–392. https://doi.org/10.1016/j.ssci.2019.02.027

Alicke, M. D., Buckingham, J., Zell, E., & Davis, T. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin, 34*(10), 1371–1381. https://doi.org/10.1177/0146167208321594

Almklov, P. G., Rosness, R., & Størkersen, K. (2014). When safety science meets the practitioners: Does safety science contribute to marginalization of practical knowledge*? Safety Science, 67*, 25–36. https://doi.org/10.1016/j.ssci.2013.08.025

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Human Error in Aviation, 37*, 101–118. https://doi.org/10.4324/9781315092898-7

Antonsen, S. (2009a). Safety culture and the issue of power. *Safety Science, 47*(2), 183–191. https://doi.org/10.1016/j.ssci.2008.02.004

Antonsen, S. (2009b). Safety culture assessment: A mission impossible*? Journal of Contingencies and Crisis Management, 17*(4), 242–254. https://doi.org/10.1111/j.1468-5973.2009.00585.x

Antonsen, S. (2009c). The relationship between culture and safety on offshore supply vessels. *Safety Science, 47*(8), 1118–1128. https://doi.org/10.1016/j.ssci.2008.12.006

Antonsen, S., Almklov, P., & Fenstad, J. (2008). Reducing the gap between procedures and practice – lessons from a succesful safety intervention. *Safety Science Monitor, 12*(1). https://www.researchgate.net/publication/229005586

Antonsen, S., Nilsen, M., & Almklov, P. G. (2017). Regulating the intangible. Searching for safety culture in the Norwegian petroleum industry. *Safety Science, 92*, 232–240. https://doi.org/10.1016/j.ssci.2016.10.013

Ashour, A., Ashcroft, D. M., & Phipps, D. L. (2021). Mind the gap: Examining work-as-imagined and work-as-done when dispensing medication in the community pharmacy setting. *Applied Ergonomics, 93*, 103372. https://doi.org/10.1016/j.apergo.2021.103372

Atak, A., & Kingma, S. (2011). Safety culture in an aircraft maintenance organisation: A view from the inside. *Safety Science, 49*(2), 268–278. https://doi.org/10.1016/j.ssci.2010.08.007

Balcerzak, T. (2017). *Scientific Journal of Silesian University of Technology . Series Transport* DOES ELECTRIC CAR PRODUCE EMISSIONS ? 94.

Berglund, J. (2020). After Fukushima : Safety culture and fostering critical thinking. *Safety Science, 124*(November 2019), 104613. https://doi.org/10.1016/j.ssci.2020.104613

Bieder, C., & Bourrier, M. (2013). *Trapping safety into rules: How desirable or avoidable is*

proceduralization? Ashgate Publishing Co. https://doi.org/10.1080/00140139.2013.819675

Biermann-Teuscher, D., Thissen, L., Horstman, K., & Meershoek, A. (2022). Safety: A Collective and Embedded Competency. An Ethnographic Study of Safety Practices at an Industrial Workplace in the Netherlands. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4094007

Blazsin, H., & Guldenmund, F. (2015). The social construction of safety: Comparing three realities. *Safety Science, 71*(Part A), 16–27. https://doi.org/10.1016/j.ssci.2014.06.001

Braithwaite, J. (2002). *Restorative Justice and Responsive Regulation*. Oxford University Press.

Brborović, O., Brborović, H., Nola, I. A., & Milošević, M. (2019). Culture of blame—an ongoing burden for doctors and patient safety. *International Journal of Environmental Research and Public Health, 16*(23). https://doi.org/10.3390/ijerph16234826

Brown, K. A., Willis, P. G., & Prussia, G. E. (2000). Predicting safe employee behavior in the steel industry: Development and test of a sociotechnical model. *Journal of Operations Management, 18*(4), 445–465. https://doi.org/10.1016/S0272-6963(00)00033-4

Bye, R. J., & Aalberg, A. L. (2020). Why do they violate the procedures? – An exploratory study within the maritime transportation industry. *Safety Science, 123*(November 2019), 104538. https://doi.org/10.1016/j.ssci.2019.104538

Carvalho, P. V. R. d., Righi, A. W., Huber, G. J., Lemos, C. de F., Jatoba, A., & Gomes, J. O. (2018). Reflections on work as done (WAD) and work as imagined (WAI) in an emergency response organization: A study on firefighters training exercises. *Applied Ergonomics, 68*, 28–41. https://doi.org/10.1016/j.apergo.2017.10.016

Catino, M., & Patriotta, G. (2013). Learning from Errors: Cognition, Emotions and Safety Culture in the Italian Air Force. *Organization Studies, 34*(4), 437–467. https://doi.org/10.1177/0170840612467156

Cavazza, N., & Serpe, A. (2009). Effects of safety climate on safety norm violations: exploring the mediating role of attitudinal ambivalence toward personal protective equipment. *Journal of Safety Research, 40*(4), 277–283. https://doi.org/10.1016/j.jsr.2009.06.002

Choudhry, R. M., Fang, D., & Mohamed, S. (2007). The nature of safety culture: A survey of the state--art. *Safety Science, 45*(10), 993–1012. https://doi.org/10.1016/j.ssci.2006.09.003

Commissie Van der Veer. (2018). Het moet en kan veiliger! Eindrapport Commissie Van der Veer.

Cooper, M. D. (2000). Towards a model of safety culture. *Safety Science, 36*(2), 111–136. https://doi.org/10.1016/S0925-7535(00)00035-7

Cox, S. J., & Cox, T. (1991). The structure of employee attitudes to safety: a European example. *Work and Stress, 5*(2), 93–106.

Cox, S. J., & Flin, R. (1998). Safety culture: Philosopher's stone or man of straw? *Work and Stress, 12*(3), 189–201. https://doi.org/10.1080/02678379808256861

Damoiseaux-Volman, B. A., Medlock, S., van der Eijk, M. D., Romijn, J. A., Abu-Hanna, A., & van der Velde, N. (2021). Falls and delirium in older inpatients: Work-as-imagined, work-as-done and preferences for clinical decision support systems. *Safety Science, 142*. https://doi.org/10.1016/j.ssci.2021.105355

Dekker, S. W. A. (2002a). Punishing people or learning from failure? The choice is ours. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Punishing+people+or+learning+from+failure?+The+choice+is+ours#0

Dekker, S. W. A. (2002b). Reconstructing human contributions to accidents: the new view on error and performance. *Journal of Safety Research, 33*(3), 371–385. https://doi.org/10.1016/S0022-4375(02)00032-4

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

Dekker, S. W. A. (2003b). When human error becomes a crime. *Human Factors and Aerospace Safety, 3*(1), 83–92.

Dekker, S. W. A. (2004). Why we need new accident models. *Human Factors and Aerospace Safety, 4*(1), pp. 1–18).

Dekker, S. W. A. (2006). The Field Guide to Understanding Human Error. Ashgate Publishing Limited.

Dekker, S. W. A. (2007a). Criminalization of medical error: Who draws the line? *ANZ Journal of Surgery, 77*(10), 831–837. https://doi.org/10.1111/j.1445-2197.2007.04253.x

Dekker, S. W. A. (2007b). Discontinuity and disaster: Gaps and the negotiation of culpability in medication delivery. *Journal of Law, Medicine and Ethics, 35*(3), 463–470. https://doi.org/10.1111/j.1748-720X.2007.00168.x

Dekker, S. W. A. (2007c). Doctors Are More Dangerous Than Gun Owners : A Rejoinder to Error Counting. *Human Factors, 49*(2), 177–184.

Dekker, S. W. A. (2007d). Eve and the serpent: A rational choice to err. *Journal of Religion and Health, 46*(4), 571–579. https://doi.org/10.1007/s10943-007-9118-1

Dekker, S. W. A. (2007e). *Just Culture; Balancing Safety and Accountability*. Ashgate Publishing Limited.

Dekker, S. W. A. (2009). Just culture: Who gets to draw the line? *Cognition, Technology and Work, 11*(3), 177–185. https://doi.org/10.1007/s10111-008-0110-7

Dekker, S. W. A. (2014a). Employees: A Problem to Control or Solution To Harness? *Professional Safety, 59*(08), 32–36.

Dekker, S. W. A. (2014b). *Safety Differently. Human factors for a new era*. Apple Academic Press Inc.

Dekker, S. W. A. (2014c). The bureaucratization of safety. *Safety Science, 70*, 348–357.

https://doi.org/10.1016/j.ssci.2014.07.015

Dekker, S. W. A. (2015). Is It 1947 Yet? https://safetydifferently.com/is-it-1947-yet/

Dekker, S. W. A. (2016). *Just culture: Restoring Trust and Accountability*. CRC Press Taylor & Francis Group.

Dekker, S. W. A. (2018). *The Safety Anarchist*. Routledge Taylor & Francis Group.

Dekker, S. W. A. (2019). *Foundations of safety science*. Taylor & Francis Group, LLC.

Dekker, S. W. A. (2020). Safety after neoliberalism. *Safety Science, 125(January),* 104630. https://doi.org/10.1016/j.ssci.2020.104630

Dekker, S. W. A., & Breakey, H. (2016). "Just culture:" Improving safety by achieving substantive, procedural and restorative justice. *Safety Science, 85*, 187–193. https://doi.org/10.1016/j.ssci.2016.01.018

Dekker, S. W. A., & Laursen, T. (2007). From Punitive Action to Confidential Reporting. *Patient Safety & Quality Healthcare, 45*(19), 1–8. http://cla.sagepub.com/cgi/doi/10.1177/0069477007310054

Dekker, S. W. A., & Pitzer, C. (2016). Examining the asymptote in safety progress: A literature Review. *International Journal of Occupational Safety and Ergonomics, 22*(1), 57–65. https://doi.org/10.1080/10803548.2015.1112104

Dekker, S. W. A., & Tooma, M. (2021). A capacity index to replace flawed incident-based metrics for worker safety. *International Labour Review, 23*(5). https://doi.org/10.1111/ilr.12210

Desai, V. M. (2010). Rule violations and organizational search: A review and extension. *International Journal of Management Reviews, 12*(2), 184–200. https://doi.org/10.1111/j.1468-2370.2008.00245.x

Edwards, J., R. D., Davey, J., & Armstrong, K. (2013). Returning to the roots of culture: A review and re-conceptualisation of safety culture. *Safety Science, 55*, 70–80. https://doi.org/10.1016/j.ssci.2013.01.004

EUROCONTROL. (2006). Establishment of "Just Culture" Principles in ATM Safety Data Reporting and Assessment - EAM2/GUI 6. In *Establishment of "Just Culture" Principles in ATM Safety Data Reporting and Assessment - EAM2/GUI 6*. (Issue March 2006). https://skybrary.aero/sites/default/files/bookshelf/235.pdf

Falconer, B. T. (2006). *Attitudes to safety and organisational culture in Australian military aviation*. Doctoral Thesis, University of New South Wales, Sydney, Australia. http://scholar.google.com.au/scholar?start=60&hl=en&num=30&cites=2395082783304815858#12

Ferraro. (2000). *Measuring Safety Climate: The Implications for Safety Performance*. Doctoral Thesis, University of Melbourne, Australia.

Filho, A. P. ., Andrade, J. C. S., & Marinho, M. M. (2010). A safety culture maturity model for petrochemical companies in Brazil. *Safety Science, 48*(5), 615–624. https://doi.org/10.1016/j.ssci.2010.01.012

Gherardi, S. (2018). A practice-Based approach to safety as an emergent competence. In *Beyond Safety Training: Embedding Safety in Professional Skills*, 11–21. https://doi.org/10.1007/978-3-319-65527-7_2

Gherardi, S., & Nicolini, D. (2002). Learning the Trade: A Culture of Safety in Practice. *Organization, 9*(2).

Glendon, A. I., & Stanton, N. A. (2000). Perspectives on safety culture. *Safety Science, 34*, 193–214.

Glesner, C., Oudheusden, M. Van, Turcanu, C., & Fallon, C. (2020). Bringing symmetry between and within safety and security cultures in high- risk organizations. *Safety Science, 132(June)*, 104950. https://doi.org/10.1016/j.ssci.2020.104950

Goodstein, J., & Aquino, K. (2010). And restorative justice for all: Redemption, forgiveness, and reintegration in organizations. *Journal of Organizational Behavior, 31*(4), 624–628. https://doi.org/10.1002/job.632

Gray, G. C. (2006). The regulation of corporate violations: Punishment, compliance, and the blurring of responsibility. *British Journal of Criminology, 46*(5), 875–892. https://doi.org/10.1093/bjc/azl005

Grote, G. (2007). Understanding and assessing safety culture through the lens of organizational management of uncertainty. *Safety Science, 45*(6), 637–652. https://doi.org/10.1016/j.ssci.2007.04.002

Guldenmund, F. W. (2000). The nature of safety culture: a review of theory and research. *Safety Science, 34*(1–3), 215–257. https://doi.org/10.1016/S0925-7535(00)00014-X

Guldenmund, F. W. (2010). *Understanding and Exploring Safety Culture*. BOXPress.

Halbesleben, J. . B., Wakefield, D. ., & Wakefield, B. . (2008). Work-arounds in health care settings: Literature review and research agenda. *Health Care Management Review,* 33(1).

Halbesleben, J. R. B., Savage, G. T., Wakefield, D. S., & Wakefield, B. J. (2010). Rework and workarounds in nurse medication administration process: Implications for work processes and patient safety. *Health Care Management Review,* 35(2).

Hale, A., & Borys, D. (2013). Working to rule, or working safely? Part 1: A state of the art review. *Safety Science, 55*, 207–221. https://doi.org/10.1016/j.ssci.2012.05.011

Hale, A. R. (1990). Safety rules o.k.?. Possibilities and limitations in behavioural safety strategies. *Journal of Occupational Accidents, 12*(1–3), 3–20. https://doi.org/10.1016/0376-6349(90)90061-Y

Hale, A. R., & Swuste, P. (1998). Safety rules: Procedural freedom or action constraint? *Safety Science, 29*(3), 163–177. https://doi.org/10.1016/S0925-7535(98)00020-4

Hayward, B. (1990). Culture, CRM and aviation safety. *Reason, 1–12*. http://www.asasi.org/papers/hayward.pdf

Health & Safety Executive. (2005). A review of safety culture and safety climate literature for the development of the safety culture inspection toolkit. Research Report 367. Health (San Francisco).

Helmreich, R. . (1998). Building Safety on the Three Cultures of Aviation. *Proceedings of the IATA Human Factors Seminar, 1999*, 39–43.

Heraghty, D., Dekker, S. W. A., & Rae, A. (2021). Modifying an accident process and its justice system – From single narratives and retribution to multiple stories and restoration. *Safety Science, 139(April 2020)*, 105248. https://doi.org/10.1016/j.ssci.2021.105248

Hollnagel, E. (2009). *The ETTO principle*. Taylor & Francis Ltd.

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. *Paper Presented at the Gestaltung Nachhaltiger Arbeitssysteme, 58. Kongress Der Gesellschaft Für Arbeitswissenschaft*, Pp.19-24.

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Hollnagel, E. (2018). *Safety-II in practice Developing the resilience potentials*. Routledge Taylor & Francis Group.

Hollnagel, E., R., W., & Braithwaite, J. (2015). *From Safety-I to Safety-II: A White Paper*.

INSAG. (1991). SAFETY SERIES No . 75-INSAG-4 INTERNATIONAL ATOMIC ENERGY AGENCY , VIENNA ,
1991 CATEGORIES IN THE IAEA SAFETY SERIES. In *Safety Series* (Issue 75).

International Civil Aviation Organization. (2018). FOSTERING JUST CULTURE IN OPERATORS AND SERVICE PROVIDERS. https://www.icao.int/APAC/Meetings/2018 APRAST12/APRAST12 WP-11 AI_5 - [NokScoot_AEROTHAI_CANSO] Just Culture.pdf#search=just culture

Iordache, V., Venera, C., & Pietreanu, B. (2016). Safety Culture in Modern Aviation Systems – Civil and Military. *INCAS Bulletin 8*(2), 135–142. https://doi.org/10.13111/2066-8201.2016.8.2.11

Johnsen, S. O., Vatn, J., Jersin, E., Veiseth, M., Rosness, R., Lamvik, G., Steiro, T., Hagen, O., & Herrera,
I. (2003). Review of existing knowledge applicable to safety culture at interfaces in European Railway undertakings. *SINTEF Report*.

Kaur, M., De Boer, R. J., Oates, A., Rafferty, J., & Dekker, S. W. A. (2019). Restorative Just Culture: a Study of the Practical and Economic Effects of Implementing Restorative Justice in an NHS Trust. *MATEC Web of Conferences, 273*(April 2018), 01007. https://doi.org/10.1051/matecconf/201927301007

Kováčová, M., Licu, A., & Bálint, J. (2019). Just Culture - Eleven Steps Implementation Methodology

for organisations in civil aviation - "JC 11." *Transportation Research Procedia, 43*, 104–112. https://doi.org/10.1016/j.trpro.2019.12.024

Krause, T. R., & Henshaw, J. L. (2005). *Leading with Safety: A Behaviour-Based Approach*. Wiley-interscience.

Lawrenson, A. J., & Braithwaite, G. R. (2018). Regulation or criminalisation: What determines legal standards of safety culture in commercial aviation? *Safety Science, 102*(July 2016), 251–262. https://doi.org/10.1016/j.ssci.2017.09.024

Lawton, R. (1998). Not working to rule: Understanding procedural violations at work. *Safety Science, 28*(2), 77–95. https://doi.org/10.1016/S0925-7535(97)00073-8

Le Coze, J. C. (2019). How safety culture can make us think. *Safety Science, 118*(December 2018), 221–229. https://doi.org/10.1016/j.ssci.2019.05.026

Lee, T. (1998). Assessment of safety culture at a nuclear reprocessing plant. *Work and Stress, 12*(3), 217–237.

Lipira, L. E., & Gallagher, T. H. (2014). Disclosure of adverse events and errors in surgical care: Challenges and strategies for improvement. *World Journal of Surgery, 38*(7), 1614–1621. https://doi.org/10.1007/s00268-014-2564-5

Lund, J., & Aarø, L. E. (2004). Accident prevention. Presentation of a model placing emphasis on human, structural and cultural factors. *Safety Science 42*(4). https://doi.org/10.1016/S0925-7535(03)00045-6

Lupton, D. (2006). Sociology and Risk. In G. M. & S. Walklate (Ed.), *Beyond the risk society* (pp. 11–24). https://doi.org/10.4324/9781315658742-2

Marx, D. (1997). Discipline: The role of rule violations. *Ground Effects: Reporting Maintenance and Groundcrew Error Reduction Efforts, 2*, 1–3. http://www.system-safety.com/articles/GroundEffects/Volume 2 Issue 4.pdf

Mascini, P. (2005). The blameworthiness of health and safety rule violations. *Law and Policy, 27*(3), 472–490. https://doi.org/10.1111/j.1467-9930.2005.00208.x

Mendoza, A. M., Liu, S. C., Dumlao, S. V., Hendricks, J. W., Son, C., Sasangohar, F., & Peres, S. C. (2020). Where Two Ends Meet: Operator and Stakeholder Perceptions of Procedures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64*(1), 1350–1354. https://doi.org/10.1177/1071181320641322

Merritt, A. C., & Helmreich, R. L. (1996). Creating and sustaining a safety culture: some practical strategies. *Paper presented at the Third Australian Aviation Psychology Symposium*, Sydney, Nov. 20-24, 1995.

Minister en Staatssecretaris van Defensie. (2018). Politieke brief bij Eindrapport Commissie Van der Veer.

Ministerie van Defensie. (2017). Kamerbrief betreffende rapport Onderzoeksraad voor Veiligheid

"Mortierongeval Mali."

Ministerie van Defensie. (2018). Een veilige defensieorganisatie. Plan van aanpak te verbetering van de veiligheid binnen Defensie.

Morris, M. W., Moore, P. C., & Sim, D. L. H. (1999). Choosing remedies after accidents: Counterfactual thoughts and the focus on fixing "human error." *Psychonomic Bulletin and Review, 6*(4), 579–585. https://doi.org/10.3758/BF03212966

National Aeronautics and Space Administration. (2003). Columbia: Accident Investigation Board. October, I(August), 100.

Neal, A., Griffin, M. A., & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behaviour. *Safety Science, 34*(1), 99–109.

Nordlöf, H., Wiitavaara, B., Winblad, U., Wijk, K., & Westerling, R. (2015). Safety culture and reasons for risk-taking at a large steel-manufacturing company: Investigating the worker perspective. *Safety Science, 73*, 126–135. https://doi.org/10.1016/j.ssci.2014.11.020

O'Keeffe, V. J., Tuckey, M. R., & Naweed, A. (2015). Whose safety? Flexible risk assessment boundaries balance nurse safety with patient care. *Safety Science, 76*, 111–120. https://doi.org/10.1016/j.ssci.2015.02.024

Obadia, I. J., Vidal, M. C. R., & e Melo, P. F. F. F. (2007). An adaptive management system for hazardous technology organizations. *Safety Science, 45*(3), 373–396. https://doi.org/10.1016/j.ssci.2006.07.002

Parker, D. L., Lawrie, M., & Hudson, P. T. W. (2006). A framework for understanding the development of organisational safety culture. *Safety Science, 44*(6), 551–562. https://doi.org/10.1016/j.ssci.2005.10.004

Parkinson, J., & Roche, D. (2004). Restorative justice: Deliberative democracy in action? Australian *Journal of Political Science, 39*(3), 505–518. https://doi.org/10.1080/103614042000295101

Patankar, M. S., Brown, J. P., Sabin, E. J., & Bigda-Peyton, T. G. (2012). *Safety Culture; Building and Sustaining a Cultural Change in Aviation and Healthcare*. Ashgate Publishing Limited.

Persson, K. D. (2012). IAEA Safety Standards on Management Systems and Safety Culture. *Atom Indonesia, 33*(1), 13–21. https://doi.org/10.17146/aij.2007.104

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate Publishing Limited.

Royal Netherlands Air Force. (2017a). Brochure 5e generatie luchtmacht.

Royal Netherlands Air Force. (2017b). Informatie sheet 5e generatie luchtmacht.

Rubin, M., Giacomini, A., Allen, R., Turner, R., & Kelly, B. (2020). Identifying safety culture and safety climate variables that predict reported risk-taking among Australian coal miners : An exploratory longitudinal study. *Safety Science, 123*(December 2019), 104564. https://doi.org/10.1016/j.ssci.2019.104564

Schein, E. H. (2004). *Organizational culture and leadership*. Jossey-Bass.

Sexton, J. B., & Klinect, J. R. (2001). The link between safety attitudes and observed performance in flight operations. In *Human Error in Aviation*, 513–518. https://doi.org/10.4324/9781315092898-25

Soeters, J. L., & Boer, P. C. (2000). Culture and flight safety in military aviation. *International Journal of Aviation Psychology, 10*(2), 111–133. https://doi.org/10.1207/S15327108IJAP1002_1

Sorensen, J. . N. (2002). Safety culture : a survey of the state-of-the-art. *Reliability Engineering and System Safety, 76*(2), 189–204.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2011*). Implementing safety management systems in aviation*. Ashgate Limited Publishing.

Taylor, F. W. (1911). *The principles of scientific management*. Harper & Brothers Publishers.

Thaden, T. L. Von, & Mitchell-Gibbons, A. (2008). The safety culture indicator scale measurement system (SCISMS). Human Factors Division Institute of Aviation. Technical Report HFD-08-03/FAA-08-2, July, 121.

The Hague: Onderzoeksraad voor de Veiligheid. (2017). Mortierongeval Mali.

Thorvaldsen, T. (2013). The importance of common sense: How Norwegian coastal fishermen deal with occupational risk. *Marine Policy, 42*, 85–90. https://doi.org/10.1016/j.marpol.2013.02.007

Tucker, A. L. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management, 22*(2), 151–169. https://doi.org/10.1016/j.jom.2003.12.006

Tucker, A. L., Zheng, S., Gardner, J. W., & Bohn, R. E. (2020). When do workarounds help or hurt patient outcomes? The moderating role of operational failures. *Journal of Operations Management, 66*(1–2), 67–90. https://doi.org/10.1002/joom.1015

Turner, K., Stapelberg, N. J. C., Sveticic, J., & Dekker, S. W. A. (2020). Inconvenient truths in suicide prevention: Why a Restorative Just Culture should be implemented alongside a Zero Suicide Framework. *Australian and New Zealand Journal of Psychiatry, 54*(6), 571–581. https://doi.org/10.1177/0004867420918659

Visitatiecommissie Defensie en Veiligheid. (2019). Tussenrapportage 2019 - Veiligheidskundige inzichten over het plan van aanpak “Een veilige defensieorganisatie.”

Visitatiecommissie Defensie en Veiligheid. (2020). Het begin is er...maar versnelling is nodig!

Visitatiecommissie Defensie en Veiligheid. (2021). Ruimte voor veiligheid. Eindrapport 2021.

Vuuren, W. Van. (2000). Cultural influences on risks and risk management: six case studies. *Safety Science, 34*, 5–9. https://docs.google.com/a/wmu.se/file/d/0B2Ifx_f4ofuXNzlSNlQ2UXV4MGs/edit

Walker, A. (2008). A qualitative investigation of the safety culture of two organisations. *Journal of*

Occupational Health and Safety - Australia and New Zealand, 24(3), 201–212.

Wang, H. (2018). Perception of safety culture : Surveying the aviation divisions of Ministry of National Defense , Taiwan , Republic of China. *Safety Science, 108*(March 2017), 104–112. https://doi.org/10.1016/j.ssci.2018.04.022

Watson, G. W., Scott, D., Bishop, J., & Turnbeaugh, T. (2005). Dimensions of interpersonal relationships and safety in the steel industry. *Journal of Business and Psychology, 19*(3), 303–318. https://doi.org/10.1007/s10869-004-2230-2

Weichbrodt, J. (2015). Safety rules as instruments for organizational control, coordination and knowledge: Implications for rules management. *Safety Science, 80*, 221–232. https://doi.org/10.1016/j.ssci.2015.07.031

Wiegmann, D. A., Zhang, H., Thaden, T. Von, Sharma, G., & Mitchell Gibbons, A. (2002). A synthesis of safety culture and safety climate research. *Aviation Research Lab, Institute of Aviation. Technical Report ARL-02-3/FAA-02-2, June*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Modality+Differences+in+Advanced+Cockpit+Displays:+Comparing+Auditory+Vision+and+Redundancy+for+Navigational+Communications+and+Traffic+Awareness#0

Woods, D. D., Dekker, S. W. A., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Ashgate Publishing Limited.

Woods, D. D., & Shattuck, L. G. (2000). Distant Supervision-Local Action Given the Potential for Surprise. *Cognition, Technology and Work, 2*(4), 242–245.

Yeoman, C. (2011). What is safety culture? Https://Sites.Google.Com/Site/Highleyyeoman/Papers-Articles/What-Is-Safety-Culture.

Zwetsloot, G. I. J. M., Kines, P., Ruotsala, R., Drupsteen, L., Merivirta, M., & Bezemer, R. A. (2017). The importance of commitment , communication , culture and learning for the implementation of the Zero Accident Vision in 27 companies in Europe. *Safety Science, 96*, 22–32. https://doi.org/10.1016/j.ssci.2017.03.001

Chapter 2 ‘Ladder’-based safety culture assessments inversely predict safety outcomes

Leonie Boskeljon-Horst, Simone Sillem, Sidney Dekker (2022)

Journal of Contingencies and Crisis Management January (1-20)

Doi: 10.1111/1468-5973.12445

‘Ladder’-based safety culture assessments inversely predict safety outcomes

Abstract

There is little empirical evidence on the predictive value of safety culture assessments (SCAs) in relation to how accident-prone an organisation might be. Recently, Antonsen not just demonstrated how a quantitative SCA mis predicted future safety outcomes, but actually showed an inverse relationship between the assessment and subsequent critical incident investigation findings. To add to our understanding, this article presents research on whether a SCA has a predictive capacity for safety outcomes. Like in Antonsen’s research, an opportunity emerged when a helicopter taxiing accident, resulting in a rotor strike occurred for a helicopter squadron that had just undergone a SCA. The assessment used ‘culture ladder’ rubrics for its findings, which allowed us to look for specific features in the subsequent independent accident investigation (in which the researchers were not involved). As with Antonsen’s findings, our research shows that a ‘ladder’-based assessment has little predictive value. Any predictive value it has is in the inverse of the assessment findings. For instance, where the SCA showed that the safety culture was very mature regarding finding a balance between safety and the mission at hand or the breaking of rules, the accident investigation pointed these out as the causes of the accident.

1. Introduction

Despite the popularity of safety culture assessments (SCAs) since the 2000s, there is little evidence for their predictive capacity for actual safety outcomes. This study asks whether a SCA offers actionable insight into how accident- or incident-prone an organisation might be. A safety culture assessment might provide a static description of the safety culture of an organisation. The same safety culture as measured with an assessment is present when an accident happens. A comparison between a proactive SCA and the reactive description of the safety culture, as found in an incident or accident investigation, should, therefore, reveal similarities. To test this assumption, and building on previous research by Antonsen (2009), a comparative study was conducted. The study investigated the results of a SCA within an operational squadron of the Royal Netherlands Air Force and compared these results to the results of an investigation of an accident that occurred within the same squadron just a few weeks after the SCA.

Antonsen has concluded that the predictive value of safety culture assessments (SCA’s) is limited (2009). In his study, a SCA using a 5-point Likert scale was compared to the results of a qualitative investigation of a major oil and gas incident. The results of a SCA using detailed workplace descriptions are compared to an accident investigation in this research article. More detail or nuance in the SCA may lead to more predictive value concerning how accident-prone a company is—whereas an assessment

using a 5-point Likert scale, which lacks detail, may not reveal that information. Therefore, the objective is to determine the predictive value of this specific assessment regarding the accident and how accident-prone this squadron was. The study reviews whether the findings of the accident investigation match the findings of the SCA. It considers whether the safety culture weaknesses found in the SCA were the causes of the incident.

In the following paragraphs, a theoretical framework is provided, and the findings are presented. First, the results of the **SCA** and the findings of the accident investigation are given, followed by a comparison of these results. This study concludes with a discussion of the findings and their practical implications.

2. Theoretical framework

2.1 Safety culture

Since the late 1980s and early 1990s, the cause of accidents has been found in the fact that operators perform their duties or interact with technology in a particular culture (Wiegmann et al., 2004), specifically a safety culture. It is commonly agreed that interest in safety culture dates from the Chernobyl accident in 1986 (Health & Safety Executive, 2005; Patankar et al., 2012; Obadia et al., 2007; Thaden & Mitchell-Gibbons, 2008; Wiegmann et al., 2002). The term 'safety culture' was first introduced by the International Nuclear Safety Advisory Group (INSAG) (1991) and defined as 'that assembly of characteristics and attitudes in organisations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance' (Obadia et al., 2007). Since the 1990s safety culture then been seen as the key to improving safety (Antonsen et al., 2017) and an enormous amount of research has been conducted to try to precisely define the concept and create a fitting model for it (Guldenmund, 2010).

2.2 The influence of safety culture on safety

Many researchers have found that safety culture has a positive impact on the safety of an organisation. For instance, Van Vuuren (2000) has found that the causes of incidents and accidents were largely due to organisational factors and that the safety culture counted for one third of those organisational factors. The majority of these safety culture failures were related to the safety attitudes of people involved in the incidents and accidents. Sexton and Klinect (2017) have found that crews with positive perceptions of safety culture could catch more errors, perform better, make fewer violations and errors and have fewer undesired aircraft states than crews with negative perceptions of safety culture. Stolzer et al. (2011) have considered safety culture, when reinforced throughout the organisation, to be one of the most effective and systemic ways to reduce the levels of accident and incident within an organisation. Berglund (2020) has found that safety culture prevented the Fukushima incident from worsening. Rubin et al. (2020) have seen safety culture as a potential cause of risk-taking behaviour, where poor safety norms led to more risk-taking behaviour.

With the Swiss cheese model (Reason, 1997) in mind, Obadia et al. (2007) have considered safety culture to be an additional defence-in-depth in the organisation, contributing to the reduction of risks at each layer of defence and, therefore, reducing the chance of an accident in the system. Zohar (2000) has found an empirical relationship between safety climate measures and the occurrence of minor work injuries. In their study, Cooper and Philips (2004) have found a link between safety climate and safety behaviour, which is probably based on perception of the importance of safety training, a factor in safety climate. Johnson (2007) has found that improving safety climate is likely to improve safe behaviour, although the results reflect a correlation and not causation.
Critiques of these findings have also been made. Recently, Antonsen (2009) has showed, based on an empirical analysis of the influence of safety culture, that it is hard to prove the relationship between safety culture and safety, but also that a proactive quantitative SCA has little predictive value regarding the occurrence of incidents and accidents. Overall, the relationship between safety culture and safety has met with considerable debate (see Henriqson et al., 2014). This research is focused on whether this concept has any predictive value in terms of safety outcomes.

2.3 Modelling safety culture

Considering safety culture as a cause of accidents requires the ability to describe and model the concept. Safety culture can be seen as something that can become better or worse as opposed to being a static concept. This view results in the maturity model concept (Corrigan et al., 2020). Westrum (2004) has distinguished the pathological, bureaucratic and generative cultures, each with a characteristic response to problems, which can be predicted based on the way information flows through the organisation (Flannery, 2001; Westrum, 2004). To depict evolutionary progression in the development of safety culture, with a true safety culture at the end of that progression (Salas et al., 2001), Hudson has elaborated on Westrum's model (1991, cited in Salas et al., 2001) by adding two stages. Hudson also changed the name bureaucratic to calculative, creating a safety culture model consisting of five stages: pathological, reactive, calculative, proactive and generative (1999). Foster and Hoult (2013 in Corrigan et al., 2020), using slightly more neutral terms, have identified five stages as vulnerable, reactive, compliant, proactive and resilient. This maturity model depicted shown in Figure 1.

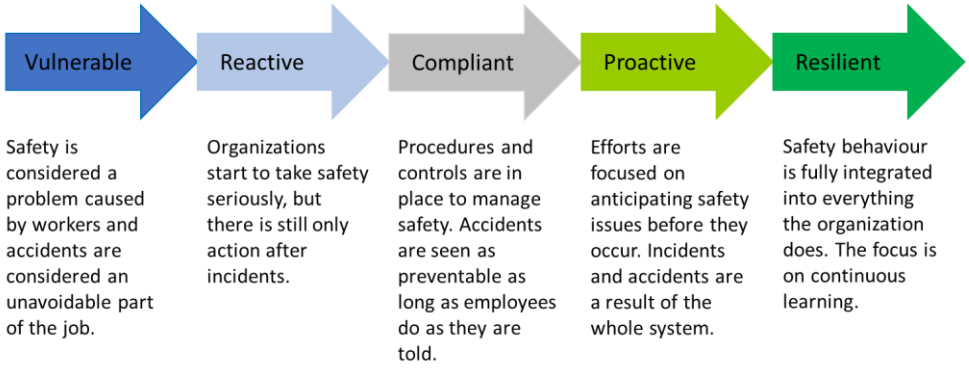

Figure 1: The maturity model of safety culture, based on (Filho et al., 2010; Hudson, 2001b; 2001c; Johnson et al,. 2003; Parker et al., 2006; Salas et al, 2001; Yeoman, 2011).

2.4 Safety culture maturity and safety

Safety culture maturity models centre on the notion that an organisation needs to develop a mature safety culture if it takes safety seriously (Hudson, 2001a). The characteristics of a mature or generative safety culture resemble those of high reliability organisations, which are known for their good safety performance records (D. L. Parker et al., 2006). To become a generative or true safety culture, an organisation must progress through the other, less mature stages (Hudson, 2001a). An organisation with an advanced, generative safety culture has fewer accidents compared to organisations with less mature safety cultures (Hudson, 2001b). A mature and effective generative safety culture can lead to substantial diminishment of the accident rate in aviation (Hudson, 2001a), whereas in organisations characterised by a pathological safety culture, the workplaces are dangerous (Hudson, 1999).

The literature on safety culture maturity claims a positive relationship between safety culture and safety, but that this relationship can be specified based on the maturity level of a given safety culture. The most mature level, generative, results in the safest performance and the least number of accidents in an organisation. However, Filho and Waterson (2018) have discussed the strengths and weaknesses of safety culture maturity models, concluding that these models assume steady progress from one level to the next, therefore, seeing values and behaviours regarding safety as static, when in fact this may not be the case. If, however, these values and behaviours are dynamic and variable, the predictive value, that is the possibility to describe the safety state of an organisation, is lost.

In this study, the relationship between the safety culture maturity concept and safety is questioned. A safety culture maturity assessment should provide insight into the protection from accidents in an organisation. If an accident happens, the investigation of this occurrence should reveal aspects of safety culture that are still underdeveloped. These should be the same aspects as revealed by the SCA, if the latter is to be considered a stable description with predictive value regarding safety and accidents. This study focuses on the predictive value of such a SCA.

3. Method

This research focused on a helicopter squadron of the Royal Netherlands Air Force (RNLAF). In 2013, at the end of a training mission, a helicopter struck a light post with its rotor blades only weeks after this squadron had taken part in a proactive SCA. This accident was investigated by an accident investigation committee. The researchers were not part of this committee. The researchers had collected SCA data using a survey a few weeks prior to the accident and were able to compare this data to the results of the accident investigation.

For the SCA a survey was developed, based on the survey used in the Hearts and Minds programme and comparable surveys by Hudson (Salas et al., 2001; Parker et al., 2006). From the literature (Zohar, 1980) (Cox & Cox, 1991) (Williamson et al., 1997) (Cox & Flin, 1998) (Cox et al., 1998) (Lee, 1998) (Cheyne & Cox, 2000) (Ferraro, 2000) (Flin et al., 2000) (Glendon & Stanton, 2000) (Griffin & Neal, 2000) (G. Grote & Künzler, 2000) (Guldenmund, 2000) (Vuuren, 2000) (Gadd & Collins, 2002) (Sorensen, 2002) (Wiegmann et al., 2002) (Patankar, 2003) (J. C. Taylor & Thomas, 2003) (Dolfini-Reed & Streicher, 2004) (GAIN Working Group E, 2004) (Gordon & Kirwan, 2004) (Hamaideh, 2004) (Cai,

2005) (Fogarty, 2005) (Health & Safety Executive, 2005) (Falconer, 2006) (Parker et al., 2006) (Thaden & Mitchell-Gibbons, 2008) (Piers et al., 2009) (Fogarty & Shaw, 2010) 129 indicators were collected. A deduction exercise, using a resemblance strategy, resulted in the reduction of these 129 indicators into nine indicators. Each indicator contains at least two, but not more than 4 items to represent this indicator, resulting in a total of 28 items. For each item, the behaviour indicative of each maturity level was described resulting in five separate situational descriptions with the same subject but different behavioural wordings (Filho et al., 2010; Hudson, 2001c; Johnsen et al., 2003). Respondents chose the description that best resembled their current working environment. Within each set, the five descriptions were in a random order, so the respondents did not know which description fit which maturity level. The statements were constructed as neutrally as possible, so it would not be clear from the wordings which statement was 'best' and which 'worst'.

In order to confirm the five descriptions per item represent accurately the five maturity stages, four other researchers received 28 envelopes, each envelope containing five cards depicting one description per card. The researchers were asked to put the descriptions in order ranging from the least developed safety culture to the highest developed safety culture. The results were discussed and a few wordings in the SCA adapted.

Using descriptions instead of a Likert scale provides a better understanding of the meaning of an answer. With a questionnaire with statements and corresponding 5 or 7-point Likert scales there is a higher risk of interpretation flaws due to misunderstanding the statement and less detail in the description.

For the assessment, specifically the mode (the answer most frequently given by the respondents), was used. When describing a concept such as safety culture, the frequency of perceptions or opinions is the starting point, not a calculated mean. This concurs with the observation that culture emerges from group processes (Kramer, 2019) and the general agreement that culture is subsequently determined by the most shared ideas, values and perceptions found in the majority of the group. This is the case, unless, of course, there is a minority with substantial formal or informal power. In that case, it is the minority that has the largest influence on culture. However, this is not the case within the squadron studied, since it is a helicopter squadron in which the principles of crew resource management are vital for both operational effectiveness and physical safety (Salas et al., 2001).

The accident investigation reported on the rotor strike. This accident was chosen for this study because of the coincidence that it happened just a few weeks after the proactive SCA. There are, luckily, few accidents a year, which makes it more difficult to find appropriate data.

A content analysis of the accident report was performed and the findings determined by the accident investigation committee were compared and related to the indicators used in the survey to determine the similarities and the possible predictive value of a SCA. Each finding in the report was related separately to a maturity description of one of the items. Relating the findings to the indicators was carried out with a 'best fit' strategy based on the description (both wording and meaning) in the report and the description of the indicators. This process allowed each finding to be related to only one indicator. For instance, the accident investigation found that the crew deviates from standard radio calls and different acknowledgements are provided instead of the prescribed ones in the training manual. This relates to the rule-making indicator at the calculative level, which states: Non-compliance to rules

and regulations is unacceptable, even if the relevance or appropriateness of the rules / regulations is doubted.

The investigation report showed no findings relating to the 8th and 9th indicators of the **survey**, role model behaviour of staff, and reinforcement or punishment. The focus of the accident investigation was on the crew and the movement of the helicopter at the time of the accident, not on commanding officers or the way safe behaviour was being reinforced. Therefore, for these indicators, a comparison cannot be made and the **SCA** results for these two indicators were omitted from this study.

The results of the SCA are presented, followed by the results of the occurrence investigation and, finally, the comparison between the two assessments is given.

4. Results of the Safety Culture Assessment (SCA)

In 2013, this squadron had 76 employees (pilots and crewmembers). Thirty-one employees returned a distributed survey, resulting in a response percentage of 41%. Although the response rate is not that high, if the intention of research is to gain in-depth knowledge about safety culture, a smaller number of respondents can still be relied on (Bergersen, 2003).

Of the 31 respondents, eleven are pilots-in-command, nine are co-pilots and eleven are loadmasters. Regarding their position within the squadron, of the 31 respondents two are senior managers within the squadron, nine middle management and the remaining 20 general operators. To guarantee the anonymity of the respondents, no other information regarding the respondents were used in this study.

Feedback provided by ten respondents after returning the survey showed that filling out the survey took on average 40 minutes due to the fact that respondents had to read a total of 240 statements (28 times 5) and had to make 28 decisions. The returned surveys showed no pattern in the answers provided, as has been seen with 5 point Likert scale questionnaires where only the third / middle category is marked. Because no pattern was detected it is assumed the respondents filled out the survey in a serious manner.

To determine if the SCA was predictive of the squadron being incident- or accident-prone, the results of the survey were compared to the findings of the accident investigation. The results of the SCA are depicted in Tables 1–7. The tables show the percentage of the respondents choosing the statement that most fitted their work environment, which was the squadron under investigation. These results show how the safety culture for the operational squadron was described. For each table, the following legend applies (Figure 2):

Safety culture maturity levels				
pathological	reactive	calculative	proactive	generative

Figure 2: Legend accompanying tables 1 to 7

Indicator 1: Rule breaking

Table 1. Indicator 1: Rule breaking

1a	Crews are trusted to recognise the situations in which compliance to regulations needs to be doubted. It is exected that crew initiatives are executed safely and that risks are mitigated by crew decision making. 63% generative	A certain amount of incompliance with rules and regulations is accepted, as long as this is discussed up front. 19% proactive	Non-compliance to rules and regulations is unacceptable, even if the relevance or appropriateness of the rules / regulations is doubted. 19% calculative	It is possible to break the rules without anyone noticing, especially when more time has passed since the last incident. 0 reactive	Regarding following the rules the saying is: "who cares, as long as were are not getting caught!. 0 pathological
1b	Breaking the rules is only allowed for safety reasons. Crews are trusted to make the right judgment. 45% proactive	On paper rules are not meant to be broken, but in practice the case is different. Management does not always live up to its word when an operational advantage is to be gained. 32% reactive	There is no such thing as breaking the rules. Each decision is about risk control. The crew is trusted to act safely. 16% generative	The most important thing is to conduct the mission, even if that means breaking rules or regulations. It is up to the crew to decide if they are capable or not. 6% pathological	It is never allowed to break the rules. The safety management system in place ensures safety. 0 calculative

The results for the first indicator showed that crews were trusted when breaking the rules for safety reasons (respectively 63% and 45%), but that it was perceived that rules were also broken for operational advantage, indicating a difference between paper and practice (32%).

Indicator 2: Empowerment of employees

Table 2. Indicator 2: Empowerment of employees.

2a	Every employee has the possibility and authority to make his/her own estimate of safety and act accordingly. 87% generative	Personnel responsible for safety has the authority to make safety decisions. 6% proactive	Only very experienced crewmembers have limited possibilities to make decisions that affect the safety of normal flight missions. 3% reactive	Pilots and crewmembers have no authority to make decisions that affect the safety of normal flight missions. 3% pathological	The authority to make safety decisions is related to someones formal positions and part of regulations. 0 calculative
2b	Crews can decide for themselves whether or not to execute risky training missions and determine if the risks are mitigated to an acceptable level. 42% generative	It is possible for employees to avoid risky training missions, however nobody does this. A mission that in hindsight should not have been executed does not result in consequences for the crew. 26% reactive	According to regulations it is not expected of employees to execute risky training missions. 19% proactive	According to regulations it is not expected of employees to execute risky training missions. 13% calculative	Even risky training missions are to be executed at all times. Flight performance is that high, this is not an issue. Refusal or avoidance is a sign of a lack of flying skills. 0 pathological
2c	Crewmembers are actively reminded of the possibility to intervene if they think safety is at risk. 53% proactive	When it comes to safety, crewmembers are equal. The authority gradient is hardly present. Copilots are actively encouraged to intervene if they deem this necessary. 22% generative	Rules, such as the two-challenge-rule, offer copilots and crewmembers, including the inexperienced ones, the possibility to contribute to the safety of the mission. 19% calculative	Crewmembers should not question actions of the captain except when these actions threaten the safety of the flight. 6% reactive	Except for total incapacitation of the captain, the co-pilot should never assume command of the aircraft. 0 pathological

The results for the second indicator indicated respondents perceived empowerment by employees relating to safety (respectively 87% and 53%). Regarding risky training missions, 42% thought they could determine themselves whether to execute them. However, 26% thought that, although possible, this was a decision never made.

Indicator 3: Balance between mission and flight safety

Table 3. Indicator 3: Balance between mission and flight safety.

3a	Safety is inseparably linked to every aspect of my job. I always make sure that my aircraft and crew are as safe as possible. 52% generative	Right after an incident has occurred I pay a little more attention to safety when doing my job. 26% reactive	Where it is possible, safety is equally important as the result when doing my job. 16% proactive	Following the rules is the most important part of my job. Safety is an expected result of compliance. 3% calculative	Personally I feel that safety issues are not the most important part of my job. 3% pathological
3b	The organization tries to make safety a top priority by acknowledging that safety can contribute to the organizational goal (operational success). The organization is pretty good in combining safety with the organizational goal, although attaining operational success is still decisive. 45% proactive	Organizational goals are and remain the most important priorities, but effort is made to comply with legal norms and regulations. 23% reactive	Most of the available time is spend on achieving operational success. Safety is guaranteed through rules and regulations and requires less attention. 19% calculative	Safety is a core activity in this organization. Safety and operational success are valued equally. Safety is a prerequisite to achieve operational success. 13% generative	Operational success is the number one priority. If this can be achieved safely it's a plus. 0 pathological
3c	Safety is part of every mission, which means that without an acceptable level of safety the mission cannot be succesful. 38% generative	Following the rules results in safety. This means no trade off between safety and performing flight missions is necessary. 19% calculative	When more time has passed since the last incident, more risks are taken during training missions and safety is taken less into account. 19% reactive	If it all comes done to it, people in this organization prefer to take safety risks to cancelling training missions. 16% pathological	When executing training missions safety receives an ever greater priority. 9% proactive

The results for the third indicator showed mixed results. Safety was seen as an inseparable part of every mission and, therefore, a priority (respectively 52%, 45% and 38%). The operational goals were considered more important and safety was usually triggered by unwanted outcomes, which diminished over time (26%, 23% and 19% respectively). Because of the rules and regulations, no tradeoff was necessary (19%).

Indicator 4: Openness

Table 4. Indicator 4: Openness

4a	I am encouraged by my leadership and co-workers to report any unsafe conditions I may observe. 46% generative	I am being encouraged to communicate my safety concerns to those who can do something about it. 27% proactive	We have a system in which I can report safety issues anonymously. All pilots and crewmembers are encouraged to use this system. 13% calculative	There is a way for me to address safety issues, however this is not encouraged. 13% reactive	Inexperienced pilots and crewmembers have no way to report safety issues. 0 pathological
4b	Sharing mistakes is SOP as far as I'm concerned. There is no shame or admiration for doing this openly. 50% generative	I am actively encouraged to openly share my mistakes, for instance in articles in our safety magazine. 40% proactive	I'd prefer my other Flight crew / air crew / flight members not to know when I made a mistake. I talk about it as little as possible. 7% reactive	I am ashamed when I make a mistake in front of my other crewmembers. I will therefore not share my mistakes with others. 3% pathological	Mistakes can be shared by reporting them anonymously in a database, which I do. 0 calculative
4c	I have no problems with writing my name on a safety report after I was unintentionally involved in an incident or accident. 48% proactive	I always report any issues that I feel are a threat to safety. 29% generative	I am willing to report below standard performance or unsafe acts of colleagues. 10% calculative	I am reluctant to report safety issues unless this can be done anonymously. 6% reactive	I am not prepared to report safety issues as I fear a negative response or personnal consequences. 6% pathological
4d	Every morning briefing starts with reporting safety issues in front of the group by the people that came across them. 42% generative	My colleagues always tell me about any safety issues they have encountered. 32% proactive	Safety issues are only discussed by the people that came across them with the people who were a witness at that time. 23% reactive	Safety issues are shared through the monthly newsletter. 3% calculative	Safety issues are not being shared among colleagues. 0 pathological

A significant proportion of respondents perceived the sharing of mistakes and safety issues as part of their job (46%, 50%, 29% and 42% respectively). They felt they were encouraged in this behaviour and saw the same behaviour in their colleagues (27%, 40%, 48% and 32% respectively).

Indicator 5: Perception of limitations

Table 5. Indicator 5: Perception of limitations

5a	When I am tired I prefer not to fly. If I have to I will use my colleagues as a safety net to compensate for the dimishing safety. 55% proactive	I am less effective when stressed or fatigued. 23% calculative	When I am tired I don't fly. The risk of fatigue is to big. There are no sufficient measures to be able to fly safely. 10% generative	When I am tired, coffe or a powernap compensates for the lack of energy so that it doesn't influences my performance. 10% reactive	Even when fatigued, I perform effectively during critical times in a flight. 3% pathological
5b	During the briefing possible emergencies and solutions are adressed. What if's, alternates, etc. are discussed to prepare the crew in case an emergency happens. 62% proactive	The crew is always prepared for emergencies with checklists, what if's, alternates, etc. It is a given that under those circumstances the quality of decision making diminishes and possible options are revised continuously during the flight. 19% generative	I am more likely to make judgment errors in an emergency. 9% calculative	My decision making ability is as good in emergencies as in routine flying conditions. 6% pathological	The decision making proces is suboptimal during emergencies, for which I can compensate with my own capacities. 3% reactive
5c	Since personnal problems can negatively affect performance, these deserve attention. It is key to discuss them up front so the rest of the crew can take this into account. 74% proactive	Personal problems can negatively affect my performance. But there are not many options to deal with this. 16% calculative	A truly professional crewmember can leave personal problems behind when flying. 6% reactive	Personnal problems are non-existent to a professional crewmember. His or her private life has nothing to do with the job. 3% pathological	Having personnal problems is a risk that is hard to mitigate and incompatible with flying. 0 generative
5d	Monitoring eachother for signs of stress or fatigue is a structural element of a pilot's or crewmember's task. 73% generative	Crewmembers should monitor each other for signs of stress or fatigue. 17% calculative	I should actively monitor my crewmembers for signs of stress or fatigue. 7% proactive	It is not necessary to monitor one another, since every pilot/crewmember is capable of doing his/her job, including when under stress or experiencing fatigue. 3% pathological	When a pilot/crewmember seems stressed or fatigued during flight, it might be wise to monitor him/her. 0 reactive

In estimating their own capabilities, the respondents saw stress, fatigue, emergencies and personal problems as possible safety threats (23%, 16% and 17%) that needed to be addressed before a flight mission (55%, 62% and 74% respectively). Preparing for emergencies and monitoring one another were seen as ways to mitigate the risk (19% and 73%).

Indicator 6: Lessons learned

Table 6. Indicator 6: Lessons learned.

6a	Nothing changes after a safety report has been written. There are no recommendations, feedback or actions taken. 39% pathological	The database containing the reported incidents and accidents is used for trend analysis. The results of this analysis are used to enhance safety. 29% proactive	Safety reports are stored in a database. On predetermined moments an analysis is made and spread throughout the organization as a form of feedback. 13% calculative	All safety reports are published. With each new report the possibility of a trend is researched. All levels of the organization actively use the the information from the safety reports in their daily work. 10% generative	Possible recommendations in a safety report only refer to the specific incident investigated. 10% reactive
6b	Incident or accident reports are spread at the supervision level after completion to share the lessons learned with alle employees. 44% proactive	Only the employees involved in an incident/accident receive feedback after an investigation. Sometimes the occurrence is mentioned in the annual report. 22% reactive	After an investigation, no lessons learned are shared. 16% pathological	Alle employees are informed about the lessons learned of an investigation, which takes the form of new regulations and an article in a safety magazine. 13% calculative	All employees are personnaly informed of an investigation and the lessons learned as a result of that investigation. 6% generative

Concerning lessons learned, 39% indicated that nothing changes, no lessons are learned, and that only employees directly involved received feedback (22%). At the same time, trend analysis was conducted (29%) and incident and accident reports were spread to share lessons learned (44%).

Indicator 7: Cause of an accident / incident

Table 7. Indicator 7: Cause of an accident / incident

7a	Accountability for an incident/accident is placed with the involved operators and their supervisors as well as the systems that failed. 35% calculative	Involved operators and their supervisors are seen as the cause of an incident/accident and are held accountable. 26% reactive	All levels in the organization are seen as possibly accountable. Management admits she should accept part of the accountability in case of an incident/accident. 23% proactive	I feel that no-one should ever be blamed for the outcome of their professional actions. Incidents/accidents are caused by the interactions in the entire system, both past and present. Management and employees share accountability. 10% generative	When an incident or accident happens it is often exclusively due to the error or violation of a single person. This individual is held accountable. 6% pathological
7b	Incidents/accidents cannot be prevented. The organization is always wary. The chance of an incident/accident can be diminshed by making safety a part of every task executed. 65% generative	Incidents/accidents are 'part of the job' and cannot be prevented. They are a fact of life. 26% pathological	Incidents/accidents are the result of system failure and lack of (compliance with) regulation. They can be prevented by focussing on implementation of systems and enough regulation. 6% calculative	Incidents/accidents are often bad luck. They can be prevented by removing the involved operators. 3% reactive	All incidents/accidents can be prevented by incorporating proactively safety in management decision making and developing proactive measures. 0 proactive
7c	In case of an incident/ accident, the investigation focusses on understanding and learning with the aim of prevention. The Why and How are more important than the Who or What. 41% generative	When investigating an incident/accident, management looks at the entire system, including processes and procedures. 25% proactive	Investigation of an incident/ accident focusses on who or what caused the occurrence (direct cause). The investigation is simple and factual. 19% reactive	In case of an incident/accident the investigation focusses on the failing system and rules that were broken 13% calculative	In case of an incident/accident the investigation focusses on finding the accountable employees. 3% pathological
7d	I see any errors that I or a colleague make as a way for the organisation to improve. The ones making a mistake, speak up about it voluntarily so everybody can learn from their experience. 54% generative	When an incident or accident has occurred this eventually always benefits safety within my organisation. 20% proactive	incidents/accidents show the organization where regulations can be improved. 20% calculative	Incidents/accidents are an unwanted disturbances of daily operations that doe not benefit the organization whatsoever. 6% pathological	Incidents and accidents show which employees are incapable of their job. 0 reactive

As can be seen in Table 7, learning and possible improvement are the focus of incident and accident investigations (41% and 54%) and the entire organisation is investigated and benefits (25% and 20%). Accountability is placed upon involved operators and their supervisors (26%) and on the systems that failed (35%). While 91% of the respondents perceived a working environment in which incidents and accidents are seen as unpreventable, 65% of the respondents thought being wary helps and 26% thought acceptance is all that is left.

5. Accident investigation findings

The accident involved a helicopter that struck a light post while taxiing. The intention was to park the helicopter after a training mission. The designated parking spot, as indicated by a marshaller, did not offer enough room to manoeuvre the helicopter. The crew feared the rotor blades would come in contact with another helicopter already parked. It was decided to park next to the already parked helicopter, but on the other side. While performing a 270-degree turn, the light post was struck. The accident investigation used a systems approach, but no specific research method was applied. The investigation report provides the findings presented in Table 8 as causally linked to the accident, which can subsequently be related to safety culture weaknesses. Table 8 represents all the factual findings written down in the report. No selection was made nor findings left out of this study. Although hindsight bias is always a risk and it is not possible to determine potential hindsight bias with the investigating committee, by using the entire investigation report the risk of hindsight bias in this study is reduced to a minimum.

Table 8. Investigation findings

No technical shortcomings found, so the investigation focuses on preparation and execution of the flight mission.
Deviation in the flight plan; supposed to be a two-ship formation, with the other PIC as section lead.
During the preparation a three hour delay occurred due to technical issues with the helicopter. The crew brought food for a proper lunch; the delay did not result in fatigue.
The flight mission lasted for 2 hours and 10 minutes and was executed without any remarkable events.
The flight mission was considered a routine flight. Crew had flown together multiple times and experienced good crew cooperation.
During the reconnaissance the crew has inspected the platform. The light posts were not mentioned by the host nation.
The reconnaissance checklist is very general, it doesn't prompt to inspect any specifics.
No written report was made of the reconnaissaince or the identified risks.
The lightposts were never seen as a risk or danger to the flight mission.
The briefing by the host nation did not mention the light posts.
Only one marshaller was to assist the parking of the helicopter. During prior parking there were two marshallers.
Crew estimates that the parking spot directed by the marshaller is to small and poses to many risks. It is considered unsafe.
Based on the directions and estimates of the loadmaster, the crew decides to park south of the already parked other helicopter.
Before the training mission, the helicopter was also parked south of the already parked other helicopter.
Copilot's attention is drawn to the marshaller who reacts very irritated upon the decision to park south and leaves the platform.
Copilot gets irritated by the irritation of the marshaller.
Physical contact between crew (loadmaster) and marshaller could have resolved the mutual misunderstanding about the parking intentions and would therefore have prevented the accident.
Crew decides to continue to taxi and park without support of the marshaller.
Since a safe parking is the responsibility of the PIC, parking without the support of a marshaller is permitted.
Crew deviates right from taxi line to provide separation with the already parked helicopter.
Subsequentely, the crew deviates left from the taxi line to provide separation with the preflight tent of which the sides started to flap.
The number one loadmaster keeps his attention at preflight tent and a group of bystanders filming the helicopter at a distance of approximately 150 m.
When the number one loadmaster can no longer see the preflight tent he directs his attention to a fire extinguishers on the platform.

Table 8 (*continued)*. Investigation findings

It is concluded that the focus of the crew was more on the consequences of the downwash during taxiing than on the dimensions of the rotorblades relating to the MOD.
The lesser focus on the dimensions of the rotorblades was a result of the fact that the week before taxiing on the platform was safe and uneventful.
At some point in the 270 degree turn the copilot notices the light post but does not consider it a risk.
Passing the light post on the right, the copilot decides not to mention it on the intercom to the rest of the crew, because it seems to fall outside the minimum obstacle distance (MOD)
The PIC notices the light post and decides to taxi straight on to create more distance with the edge of the platform.
There is no communication within the crew regarding the decision to taxi straight on.
At the edge of the platform the PIC decides to deviate left of the taxi line to create space for the tail of the helicopter. The focus is on the edge of the platform, not on the light post.
The number one loadmaster has not seen the light post.
The PIC is relatively unexperienced, less than 60 flight hours as a PIC.
The number 2 loadmaster is relatively unexperienced, just over 90 flight hours of experience.
All crewmembers were fit to fly. They had had enough rest and no personal problems.
The RNLAF has to be compliant with the militaire aviation requirements.
RNLAF rules and regulations ensure this compliance.
Every employee is expected to know the relevant rules and regulations, so that they will be compliant.
During training of pilots and loadmasters enough attention is given to the subject of distance between rotorblades and obstacles.
The crew was distracted due to the vehicle parked behind the already parked helicopter, the behaviour of the marshaller, the tent on the platform and the group of bystanders.
As a result of the cognitive distraction the situation awareness of the crew dimished. Human attention span is limited.
Based on the lookout sections of the PIC, COP and number one loadmaster the lightpost could have been identified as a risk to safety had they realised that the light post was within the MOD.
According to the training manual a distance of a minimum of 5 m. has to be guaranteed.
Based on the CVR it is concluded that the crew deviates from standard radio calls. Different acknowledgements are provided instead of the prescribed ones in the training manual.
A turn to the left is executed without clearance from the loadmaster.
In flight training issues such as platforms that deviate from minimum separation distances are not addressed, which is considerd a training deficiency.
The crew experienced a false sense of safety while taxiing on a platform that does not entail the minimum separation distances for helicopters.
A similar accident, in which a helicopter collided with a container while taxiing, happened three years prior to this accident.

6. Comparing the safety culture assessment to the accident investigation

Based on a content analysis and best fit strategy using resemblance, the findings in the investigation report were related to the indicators used in the survey and the maturity level according to models by Hudson (Salas et al., 2001) and Foster and Hoult (Foster & Hoult, 2013) were determined. The results are shown in Tables 9–15. Eleven findings in the investigation report could not be related to the indicators because they either displayed facts unrelated to the crew, such as the host nation, or content not covered by the indicators. These findings were, therefore, not compared.

Indicator 1: Rule breaking

Table 9. Comparison occurrence results to SCA indicator 1 Rule braking.

Findings accident investigation	indicator	level
The RNLAF has to be compliant with the militaire aviation requirements.	rule breaking	calculative
RNLAF rules and regulations ensure this compliance.	rule breaking	calculative
Every employee is expected to know the relevant rules and regulations, so that they will be compliant.	rule breaking	calculative
According to the training manual a distance of a minimum of 5 m. has to be guaranteed.	rule breaking	calculative
Based on the CVR it is concluded that the crew deviates from standard radio calls. Different acknowledgements are provided instead of the prescribed ones in the training manual.	rule breaking	calculative
A turn to the left is executed without clearance from the loadmaster.	rule breaking	calculative

Regarding the first indicator, rule breaking, the SCA results showed that the first item was perceived at the generative level (63%), the second item at the proactive (45%) and the reactive level (32%). The accident investigation, however, showed fi

ndings that all related to the calculative level, indicating that the results of the SCA do not match the results of the accident investigation. The SCA describes a safety culture that acknowledges that to be and stay safe, rules sometimes need to be broken. It describes a safety culture in which the pilots and crewmembers know the difference between situations in which rules can be broken for safety reasons and situations in which compliance with the rules is necessary. The accident investigation, however, showed that, in this case, the crew could not see the need to comply. The crew deviated a few times from rules and procedures, which, according to the investigation report, led to the accident.

Indicator 2: Empowerment of employees

Table 10. Comparison occurrence results to SCA indicator 2 Empowerment of employees.

Findings accident investigation	indicator	level
Crew estimates that the parking spot directed by the marshaller is to small and poses to many risks. It is considered unsafe.	empowerment of employees	generative
Based on the directions and estimates of the loadmaster, the crew decides to park south of the already parked other helicopter.	empowerment of employees	generative
Crew decides to continue to taxi and park without support of the marshaller.	empowerment of employees	generative
Since a safe parking is the responsibility of the PIC, parking without the support of a marshaller is permitted.	empowerment of employees	proactive
Crew deviates right from taxi line to provide separation with the already parked helicopter.	empowerment of employees	generative
Subsequentely, the crew deviates left from the taxi line to provide separation with the preflight tent of which the sides started to flap.	empowerment of employees	generative
At some point in the 270 degree turn the copilot notices the light post but does not consider it a risk.	empowerment of employees	generative
The PIC notices the light post and decides to taxi straight on to create more distance with the edge of the platform.	empowerment of employees	generative
At the edge of the platform the PIC decides to deviate left of the taxi line to create space for the tail of the helicopter. The focus is on the edge of the platform, not on the light post.	empowerment of employees	generative

Regarding the second indicator, empowerment of employees, SCA results showed the respondents perceived the first item at the generative level (87%), the second item at both a generative (42%) and reactive (26%) level (pertaining to the avoidance or risky missions, which was not the case in this investigated situation) and the third item at the proactive level (53%). The accident investigation showed all generative results except for the fourth result, which corresponds to the proactive level. It can be seen from the investigation results that the investigated behaviour showed empowerment in practice,

as predicted in theory by respondents. This means that the behaviour shown matched the behaviour described in the SCA.

Indicator 3: Balance between mission and flight safety

Table 11. Comparison occurrence results to SCA indicator 3 Balance between mission and flight safety.

Findings accident investigation	indicator	level
During the reconnaissance the crew has inspected the platform. The light posts were not mentioned by the host nation.	balance mision and flight safety	calculative
No written report was made of the reconnaissance or the identified risks.	balance mision and flight safety	calculative
The lightposts were never seen as a risk or danger to the flight mission.	balance mision and flight safety	reactive

Regarding the third indicator, balance between mission and flight safety, the SCA results showed the respondents perceived the first item at the generative level (52%), the second item at the proactive (45%) and reactive (23%) level and the third item at the generative (38%) and calculative (19%) level. The accident investigation, however, showed results that corresponded with either the calculative or reactive level, indicating that the SCA showed a more positive perception of the way safety was incorporated in the daily work than the accident investigation. Although, in the SCA, safety was seen as inextricably linked to the mission by respectively 52%, 45% and 23%, the accident investigation showed different outcomes. The SCA and the accident investigation described different behaviour and, therefore, different safety cultures.

Indicator 4: Openness

Table 12. Comparison occurrence results to SCA indicator 4 Openness.

Table 12: comparision occurrence results to SCA indicator 4 Openness		
Findings accident investigation	**indicator**	**level**
The flight mission was considered a routine flight. Crew had flown together multiple times and experienced good crew cooperation.	openness	proactive
Passing the light post on the right, the copilot decides not to mention it on the intercom to the rest of the crew, because it seems to fall outside the minimum obstacle distance (MOD)	openness	calculative
There is no communication within the crew regarding the decision to taxi straight on.	openness	pathological

Regarding the fourth indicator, openness, the SCA results showed the respondents perceived the first item at both the generative and proactive level (46% and 27% respectively), the second item at the generative level (50%), the third item at the proactive (48%) and generative (29%) level and fourth item again at the generative and proactive level (42% and 32% respectively). As concluded from the SCA, openness was part of their daily existence, according to the respondents. The results in the accident investigation corresponded with the pathological, calculative and proactive level. Although a positive climate resulting from good crew cooperation was noticed, the lack of communication was seen as a lack of openness and causally related to the accident. However, based on the investigation report, it is unclear whether the lack of communication was due to a lack of openness or simply the result of not seeing the need to communicate. Based on the description in the investigation report, the safety culture outlined in the investigation report does not match the safety culture detailed in the SCA.

Indicator 5: Perception of limitations

Table 13. Comparison occurrence results to SCA indicator 5 Perception of limitations.

Findings accident investigation	indicator	level
During the preparation a three hour delay occurred due to technical issues with the helicopter. The crew brought food for a proper lunch; the delay did not result in fatigue.	perception of one's own limitations	generative
Copilot's attention is drawn to the marshaller who reacts very irritated upon the decision to park south and leaves the platform.	perception of one's own limitations	pathological
Copilot gets irritated by the irritation of the marshaller.	perception of one's own limitations	pathological
The number one loadmaster keeps his attention at preflight tent and a group of bystanders filming the helicopter at a distance of approximately 150 m.	perception of one's own limitations	pathological
When the number one loadmaster can no longer see the preflight tent he directs his attention to a fire extinguishers on the platform.	perception of one's own limitations	pathological
It is concluded that the focus of the crew was more on the consequences of the downwash during taxiing than on the dimensions of the rotorblades relating to the MOD.	perception of one's own limitations	pathological
The lesser focus on the dimensions of the rotorblades was a result of the fact that the week before taxiing on the platform was safe and uneventful.	perception of one's own limitations	pathological
All crewmembers were fit to fly. They had had enough rest and no personal problems.	perception of one's own limitations	generative
The crew was distracted due to the vehicle parked behind the already parked helicopter, the behaviour of the marshaller, the tent on the platform and the group of bystanders.	perception of one's own limitations	pathological
As a result of the cognitive distraction the situation awareness of the crew dimished. Human attention span is limited.	perception of one's own limitations	pathological
The crew experienced a false sense of safety while taxiing on a platform that does not entail the minimum separation distances for helicopters.	perception of one's own limitations	pathological

Regarding the fifth indicator, perception of one's own limitations, the SCA results showed the respondents perceived the first three items at the proactive level (55%, 62% and 74% respectively) and the fourth item at the generative level (73%), indicating that personal limitations needed to be addressed, mitigated and monitored. Of course, occurrences, such as being annoyed or distracted, are relatively common. The results of the SCA did not indicate that these emotions would no longer occur, but that they were acknowledged and mitigated to ensure a safe flight. The accident investigation showed nine findings that corresponded to the pathological level and two findings corresponding with the generative level. Before the flight, the crew had taken care of their physical wellbeing (rest, food). However, during the flight, the taxiing phase, the crew realised they were irritated and distracted, but did not counter these feelings, discuss them or otherwise mitigate them to ensure the safety of this phase. The investigation report showed they failed to understand the effects of their emotional wellbeing. The safety culture described by the accident investigation did not match the safety culture described by the SCA.

Indicator 6: Lessons learned

Table 14. Comparison occurrence results to SCA indicator 6 Lessons learned

Findings accident investigation	indicator	level
A similar accident, in which a helicopter collided with a container while taxiing, happened three years prior to this accident.	Lessons learned	reactive

Regarding the sixth indicator, lessons learned, the SCA results showed the respondents perceived the first item at the pathological (39%) and proactive (29%) level and the second item at the proactive (44%) and reactive (22%) level. The accident investigation revealed only one finding corresponding to this indicator, which was at the reactive level. Although the SCA seemed to describe an organisation that learned from previous incidents and accidents, it also showed that it did not. The results of the SCA were ambiguous. The accident investigation concluded that the organisation did not learn from previous

occurrences, given the similarities with a previous accident. The safety culture described by the accident investigation, therefore, only partly matched the safety culture described by the SCA.

Indicator 7: Cause of an accident / incident

Table 15. Comparison occurrence results to SCA indicator 7 Cause of an accicent/incident

Findings accident investigation	indicator	level
No technical shortcomings found, so the investigation focuses on preparation and execution of the flight mission.	cause of an accident/incident	reactive
Physical contact between crew (loadmaster) and marshaller could have resolved the mutual misunderstanding about the parking intentions and would therefore have prevented the accident.	cause of an accident/incident	calculative
Based on the lookout sections of the PIC, COP and number one loadmaster the lightpost could have been identified as a risk to safety had they realised that the light post was within the MOD.	cause of an accident/incident	calculative

Regarding the seventh indicator, cause of an accident or incident, the SCA results showed respondents perceived the first item to be at the calculative (35%) and reactive (26%) level, the second and fourth item at the generative level (65% and 54% respectively) and the third item at the generative (41%) and proactive (25%) level. The accident investigation showed findings corresponding to the reactive and calculative level. This indicates that the assessment showed a safety culture in which accidents and incidents were seen as emerging from the entire organisation and something the organisation could learn from, but that responsibility and accountability were often placed on the involved employees. The investigation report partly reflected the same safety culture. Regarding responsibility and accountability, the accident was seen as preventable had the involved operators been more careful.

Although this comparison seems to show differences and similarities between the two assessments (SCA and accident investigation), it is not clear how significant the differences are or how well the results match. The results of both assessments are, therefore, depicted in graphs, to study the differences further. For each graph, the horizontal axis depicts the number of items in the survey and the number of findings in the accident investigation. The vertical axis depicts the maturity level ranging from 1–5, with 1 indicating the pathological level and 5 the generative level. The orange bars represent the results from the SCA and the blue bars the results from the accident investigation. The individual survey items are not compared to the individual accident investigation findings, only the pattern found in the SCA is compared to the pattern found in the accident investigation.

When comparing the results of the SCA and the accident investigation, the same pattern is found in Indicators 3 - balance between mission and flight safety, 4 - openness, 6 - lessons learned and 7 - cause of an accident or incident. The following four graphs show this pattern.

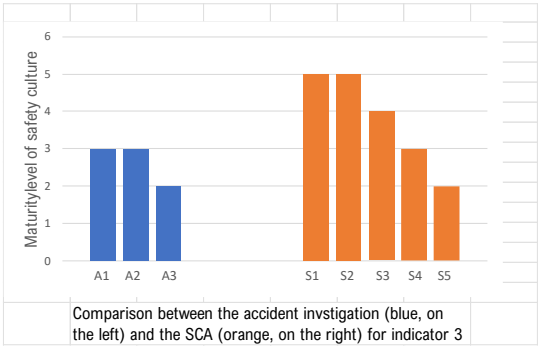

Figure 3. Indicator 3: Balance between safety and flight mission.

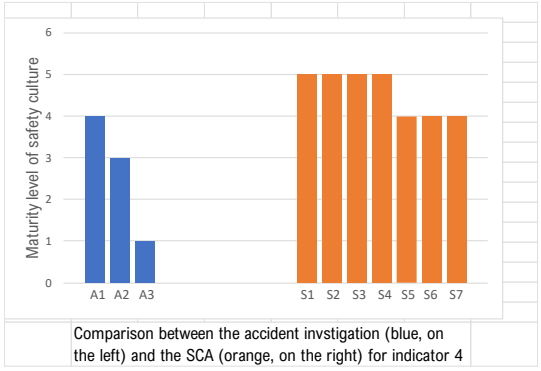

Figure 4. Indicator 4: Openness.

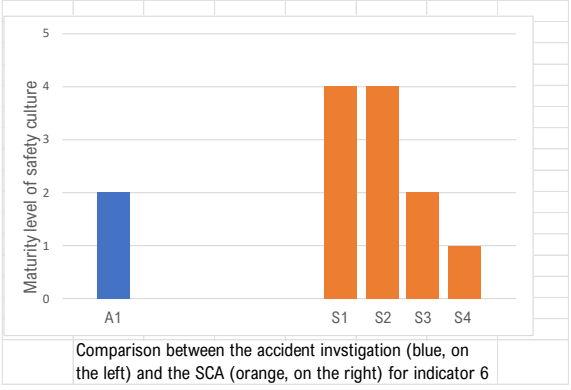

Figure 5. Indicator 6: Lessons learned.

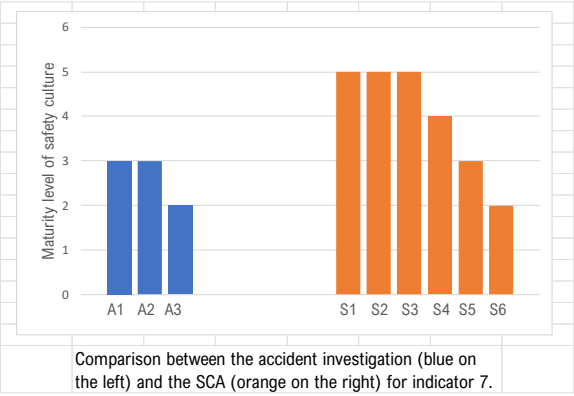

Figure 6. Indicator 7: Cause of incident/accident

Based on these graphs, the SCA shows a more positive, more mature safety culture for these four indicators than the accident investigation. These indicators relate to the role and behaviour of the operators. According to the SCA the operators working at the involved helicopter squadron maintained a good balance between flight mission and safety. The operators were open, learned from previous incidents or accidents and considered an incident or accident to be a result of the entire system, although they recognised that the operators involved were marked as responsible. The accident investigation, however, revealed the operators as the cause of the helicopter accident, supported by a lack of openness, not enough priority placed on safety and the fact that a similar accident happened three years previously. For these four indicators, the operators had a more positive perception of their safety culture than the organisational investigation committee.

The results for Indicator 1 - rule breaking show the same pattern when looking at the content of the indicator.

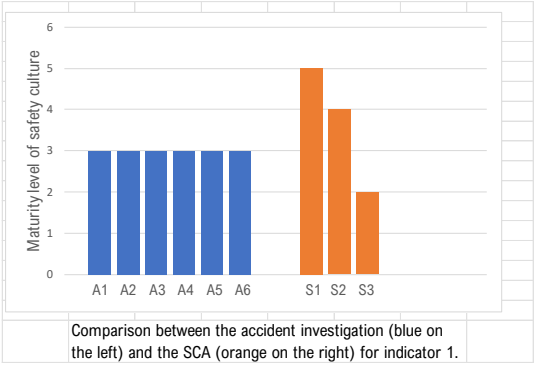

Figure 7. Indicator 1: Rule breaking

Although the maturity levels suggest a linear development for the subject of rule breaking, it is more of a parabola development. The underlying theme regarding the breaking of rules is the amount of freedom one experiences. According to the maturity ladder for both the pathological and the generative level, operators can break the rules if they think they need to. In a pathological safety culture, this is possible

because succeeding in the mission is the most important goal and breaking the rules is allowed to achieve this goal. In a generative safety culture, this is possible because operators are trusted to

Figure 8. Scoring of indicator 1: Rule breaking

recognise when they need to comply and when they need to deviate regarding safety and risk control. In both the pathological and generative description, operators have considerable freedom regarding breaking the rules. This leads to a different scoring, with the pathological and generative levels receiving similar codes, as do the reactive and proactive levels, based on the amount of freedom. This line of thinking results in the pathological and generative levels being represented by a 5, the reactive and proactive levels by a 3 and the calculative levels by a 1. The numbers 5, 3 and 1 are arbitrary, chosen to show the difference more clearly than in the case of the numbers 3, 2 and 1. In this parabola graph, the numbers 1–5 on the horizontal axis represent the five maturity levels, which are pathological (1), reactive (2), calculative (3), proactive (4) and generative (5). The vertical axis represents the scoring of the levels. This shows the parabola effect of this particular indicator.

Translating the parabola effect to the SCA results and investigation findings results in a different depiction of the original graph of indicator 1:

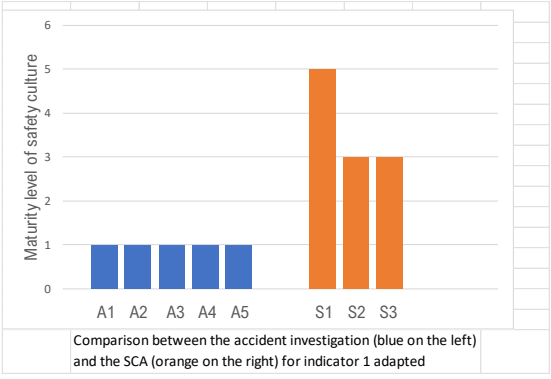

Figure 9. Indicator 1: Rule breaking adapted.

Taking into account the parabola effect, it shows, as with Indicators 3, 4, 6 and 7, that for Indicator 1 the operators perceived their safety culture as more positive and more mature than in the investigation

committee findings. According to the investigation committee, based on the military flight regulations, which are also part of the safety culture, operators may not break the rules. According to the operators, however, they work in a safety culture that does allow them to break the rules when necessary.

Indicator 2 - empowerment of employees seems to show a different pattern, in which the accident investigation reveals a more positive and more mature safety culture than the SCA.

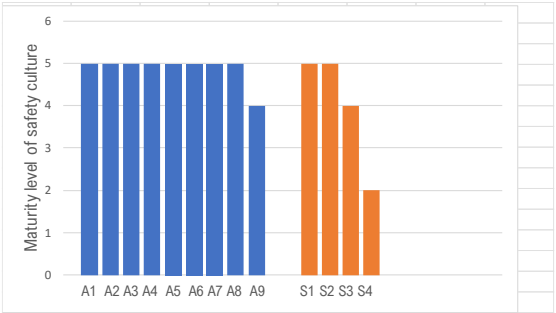

Figure 10: Indicator 2: Empowerment of emplouees.

What leads to this pattern is the perception of operators, seen in the SCA, that risky training missions can be avoided but would not be, indicating a focus more on mission and goal completion than on safety. In practice, however, as seen by the accident investigation, parking at the parking spot indicated by the marshaller, the final part of the mission, was refused in favour of a parking spot deemed to be safer. The operators did not predict their own behaviour regarding the use of their judgement to make a decision. For the investigation results, a relationship can be seen with Indicator 7 - cause of an incident or accident. According to the investigation, the cause of this accident was the involved crew, considering the decisions they made, which they should not have. In the investigation too much empowerment (freedom) was found, resulting in the accident.

The fifth indicator, perception of limitations, also seems to convey a mixed message.

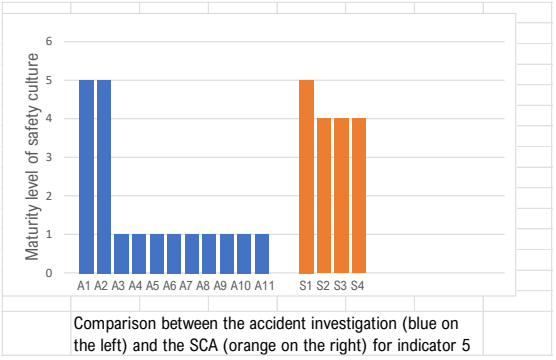

Figure 11. Indicator 5: Perception of limitations

When looking closer at the content of both the SCA and the accident investigation, it becomes clear that the operators perceive their safety culture as more positive and mature than was found during the accident investigation. The difference between the two was significant. According to the investigation, operators showed a mature safety culture when preparing for a flight, as evidenced in their buying lunch and being well rested, but that, during the flight, the crew displayed behaviour resembling a rather immature safety culture.

Looking at the seven indicators together, the operators considered their safety culture rather mature and indicated that they played an active role in achieving a safe flight. The investigation showed a more immature safety culture in which the behaviour of the operators led to a lack of safety rather than safety. The described safety cultures do not match and, therefore, the SCA seems to have limited predictive value regarding the situation found in the accident investigation.

7. Discussion

The findings of this study show, as did the research by Antonsen (2009), that the results of the accident investigation do not match the results of the SCA. The comparison of the SCA with the accident investigation shows that the two different methods describe two different safety cultures, adding to the literature by demonstrating the limitative predictive value of a SCA.

Based on the SCA, it seemed very unlikely an accident would happen that was caused by factors, such as empowerment or perception of limitations. Yet this is exactly what was determined by the accident investigation committee. The occurrence of an accident is not remarkable. There is always a chance an accident happens. However, according to the theoretical framework, the chance of an accident is diminished in the mature safety culture which the SCA described (Hudson, 1999; Hudson, 2001c). What is remarkable, however, is that the accident investigation determined factors reflecting a mature safety culture, such as empowerment, as the cause of this specific accident. This means that the SCA has only limited predictive value. A safety culture maturity description of a workplace or an organisation does not provide protection from accidents. It might in fact provide an illusion, even leading to the idea of a 'check in the box' once a SCA shows an organisation has reached the proactive or even the generative maturity level.

This finding is supported by the fact that the SCA and the accident investigation were only a few weeks apart. The SCA was conducted in September. The accident happened in October of the same year. No significant changes occurred within the squadron organisation during that time. Safety culture is seen as a derivative of culture (Guldenmund, 2010), which is seen as a stable concept (Schein, 2004). It is, therefore, highly unlikely that the safety culture of the squadron changed between the assessment and the accident.

Perhaps the difference in results when comparing the two measurements was caused by the instruments used. However, the SCA is based on a validated and well-used instrument (Energy Institute, 2015, 2019). The difference between the original instrument and the SCA is in the wordings used, matching the RNLAF organisation. The underlying theoretical framework is the same. The accident investigation was the official instrument to conduct an accident investigation, based on and driven by

international norms, agreements and standards applicable to air safety investigations (International Civil Aviation Organisation, 1993). The investigation committee followed these and internal guidelines and procedures regarding the process of accident investigation.

A limitation in this study was the response rate of 41% of the SCA, although previous research (Bergersen, 2003) shows that the number of respondents is adequate for scientific analysis. Furthermore, the balance between the number of pilots-in-command and co-pilots within the response group resembles the balance in the squadron. Only the loadmasters are a little under represented. The balance regarding position (senior management, middle management, operators) again resembles the squadron. These representations add to the generalizability of the findings of the SCA. Another limitation could be the content of the SCA. The SCA is based on a literature study of literature on safety culture indicators, which can never be complete, inevitably resulting in questions not asked. This research has been set as much as possible in the context of available scholarship, as described in the method section.

The question remains of what to do with two conflicting descriptions of the supposedly same safety culture. According to Antonsen (2009), it makes sense to consider the description based on the accident investigation as most accurate, since investigations can be seen as more thorough and are considered more important in the organisation. The reason for this preference is that an organisation is highly motivated to find the cause of an incident or accident in order to prevent future ones. That motivation is lacking when 'simply' assessing the safety culture when nothing is out of the ordinary. When conducting an accident investigation, the investigators describe the behaviour of others. In principle, they have nothing to gain by describing this behaviour more positively or negatively than observed. The SCA asks respondents to describe their workplace and their own behaviour in it, as do most SCAs. It is possible that the respondents have a more positive perception of this than is supported by reality. Unfortunately, the choice to have respondents describe their safety culture instead of having an external observer make the evaluation falls outside the scope of this study.

The choice to consider the squadron safety culture description based on the accident investigation as the most accurate leads to the conclusion that the predictive value of the SCA is limited but that it inversely predicted the safety outcomes. Aspects of the safety culture that, according to the SCA, are very mature (proactive and generative) were identified as the causes of this accident.

The instruments used, the SCA and the accident investigation, have a different focus. The SCA focuses, as do all safety culture maturity assessment instruments, on the way safety is created or enhanced (Parker et al., 2006). The results show the perceived safety culture in the absence of an accident. The results of the accident investigation show the accident investigators' perception of that same safety culture in the presence of an accident, that is, in the absence of safety, according to the accident investigators.

Perhaps it is this difference in research focus or aim that makes it impossible to compare the content and even find an inverse prediction. If prediction is what an organisation aspires to, it might be more fruitful to design an instrument that measures safety and to subsequently test this instrument for predictive value not only in situations characterised by a lack of safety but also in situations where safety was actually achieved. A safety culture maturity assessment apparently cannot accomplish that.

This is supported by the literature on safety culture as a complex and dynamic concept which assessments using surveys are unable to convey (Guldenmund, 2007) and a quantitative approach is

limited when the aim is prevention (Guldenmund, 2000). Safety culture is continually changing as a result of the dynamic social reality surrounding it (Richter & Koch, 2004) (Gephart et al., 2009) (Silbey, 2009). Furthermore, accident and incidents cannot be seen as a simple indicator of safety culture (Cooper, 2000)(Richter & Koch, 2004).

Based on this, one could ask the question if there is an instrument that does adequately provide insight into the safety culture of an organization, be it an assessment or an accident investigation tool. As the literature shows, as mentioned above, questionnaires and assessments don't do the complex phenomenon justice. Neither do most accident investigation methods as they focus on what went wrong, which is usually the exception (Hollnagel, 2014). What could be done however, is to study the way people do their work. Within a complex setting, operators are faced with multiple goals that are often not compatible with one another (Woods et al., 2010). Operators find ways to deal with these goal conflicts by employing what is called local ingenuity (Boskeljon-Horst et al., 2022): routines that help resolve goal conflicts and become part of the regular repertoire of operators, not necessarily contrary to the literal wording or intent of rules and procedures, but were not originally intended and not included in current documentation. How local ingenuity is employed to achieve safety, one of goals to be achieved, adds to the understanding of the workings of a complex system operators are part of and provides more insight than a SCA or an accident investigation.

8. Conclusion

The research aim in this article was to determine whether the SCA possesses predictive value regarding how accident-prone an organisation is. In the case under study, the organisation was a helicopter squadron. However, the official instrument to conduct an accident investigation had different results and a different safety culture description than the safety culture maturity assessment. The results of the comparison show that the SCA has limited predictive value for the safety culture during an accident but that it inversely predicts the safety outcome. Factors indicative of a mature, proactive and generative, safety culture were found to be the causes of this particular accident. The findings by Antonsen (2009), based on the comparison of a quantitative SCA to an accident report, were replicated, this time using a SCA with detailed workplace descriptions. This research reinforces the conclusion Antonsen draws regarding the limited predictive value of safety culture maturity assessments. Development of a means to predict safety outcomes is planned using micro experiments within a helicopter squadron. These results are discussed elsewhere.

References

Antonsen, S. (2009). Safety culture assessment: A mission impossible? *Journal of Contingencies and Crisis Management*, *17*(4), 242–254. https://doi.org/10.1111/j.1468-5973.2009.00585.x

Antonsen, S., Nilsen, M., & Almklov, P. G. (2017). Regulating the intangible. Searching for safety culture in the Norwegian petroleum industry. *Safety Science*, *92*, 232–240. https://doi.org/10.1016/j.ssci.2016.10.013

Bergersen, C. (2003). Tool to be used to survey and improve safety culture in the European railway industry. *Unpublished Master's Thesis, Norwegian University of Science and Technology*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Tool+To+Be+Used+To+Survey+and+Improve+Safety+Culture+In+The+European+Railway+Industry#0

Berglund, J. (2020). After Fukushima : Safety culture and fostering critical thinking. *Safety Science*, *124*(November 2019), 104613. https://doi.org/10.1016/j.ssci.2020.104613

Boskeljon-Horst, L., Boer, R. J. De, Sillem, S., & Dekker, S. W. A. (2022). Goal Conflicts , Classical Management and Constructivism : How Operators Get Things Done. *Safety*, *8*(37).

Cai, W. (2005). *The impact of safety culture on performance: a case study of a construction company*. Indiana University.

Cheyne, A. J. ., & Cox, S. . (2000). Assessing safety culture in offshore environments. *Safety Science*, *34*, 111–129.

Cooper, M. D. (2000). Towards a model of safety culture. *Safety Science*, *36*(2), 111–136. https://doi.org/10.1016/S0925-7535(00)00035-7

Cooper, M. D., & Phillips, R. A. (2004). *Exploratory analysis of the safety climate and safety behavior relationship*. *35*, 497–512. https://doi.org/10.1016/j.jsr.2004.08.004

Corrigan, S., Kay, A., Ryan, M., Brazil, B., & Ward, M. E. (2020). *Human factors & safety culture : Challenges & opportunities for the port environment*. *125*(July 2017). https://doi.org/10.1016/j.ssci.2018.02.030

Cox, S. J., & Cox, T. (1991). The structure of employee attitudes to safety: a European example. *Work and Stress*, *5*(2), 93–106.

Cox, S. J., & Flin, R. (1998). Safety culture: Philosopher's stone or man of straw? *Work and Stress*, *12*(3), 189–201. https://doi.org/10.1080/02678379808256861

Cox, S. J., Tomás, J. M., Cheyne, A., & Oliver, A. (1998). Safety culture: The prediction of commitment to safety in the manufacturing industry. *British Journal of Management*, *9*(SPEC. ISS.). https://doi.org/10.1111/1467-8551.9.s1.2

Dolfini-Reed, M. A., & Streicher, B. L. (2004). Creating a safety culture. *CRM D001056.A2/1Rev*. https://doi.org/10.21236/ada605037

Energy Institute. (2015). *Improving safety culture through Hearts and Minds*. *September*. https://minerva.jrc.ec.europa.eu/EN/content/minerva/7592dbf6-056b-46e6-96c0-2e580b9c3b69/201509mjv_safety_culture_nlkinglawrie

Energy Institute. (2019). Understanding your HSE culture. *Hearts & Minds*. https://publishing.energyinst.org/heartsandminds/toolkit/UYC

Falconer, B. T. (2006). Attitudes to safety and organisational culture in Australian military aviation. *Aviation*, 337–350. http://scholar.google.com.au/scholar?start=60&hl=en&num=30&cites=2395082783304815858#12

Ferraro. (2000). *Measuring Safety Climate: The Implications for Safety Performance* (p. 1217).

Filho, A. P. ., Andrade, J. C. S., & Marinho, M. M. (2010). A safety culture maturity model for petrochemical companies in Brazil. *Safety Science*, *48*(5), 615–624. https://doi.org/10.1016/j.ssci.2010.01.012

Filho, A. P. G., & Waterson, P. (2018). Maturity models and safety culture: A critical review. *Safety Science*, *105*(June), 192–211. https://doi.org/10.1016/j.ssci.2018.02.017

Flannery, J. A. (2001). Safety Culture and its measurement in aviation. *Transport*, *36*(25), 1–29. http://www.mendeley.com/research/safety-culture-measurement-aviation-1/

Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science*, *34*(1–3), 177–192. https://doi.org/10.1016/S0925-7535(00)00012-6

Fogarty, G. J. (2005). Psychological strain mediates the impact of safety climate on maintenance errors. *International Journal of Applied Aviation Studies*, *5*(1), 53–63.

Fogarty, G. J., & Shaw, A. (2010). Safety climate and the Theory of Planned Behavior: Towards the prediction of unsafe behavior. *Accident Analysis and Prevention*, *42*(5), 1455–1459. https://doi.org/10.1016/j.aap.2009.08.008

Foster, P., & Hoult, S. (2013). *The Safety Journey: Using a Safety Maturity Model for Safety Planning and Assurance in the UK Coal Mining Industry*. 59–72. https://doi.org/10.3390/min3010059

Gadd, S., & Collins, A. M. (2002). *Safety Culture: A review of the literature, Report No. HSL/2002/25*.

GAIN Working Group E. (2004). *A roadmap to a just culture: enhancing the safety environment*.

Gephart, R. P., Van Maanen, J., & Oberlechner, T. (2009). Organizations and risk in late modernity. *Organization Studies*, *30*(2–3), 141–155. https://doi.org/10.1177/0170840608101474

Glendon, A. I., & Stanton, N. A. (2000). Perspectives on safety culture. *Safety Science*, *34*, 193–214.

Gordon, R., & Kirwan, B. (2004). Developing a safety culture in a research and development environment: Air Traffic Management domain. *Europe Chapter of the Human Factor and Ergonomic Society Conference*, *129*(August 2014), 2865. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.79.9628&rep=rep1&type=pdf

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: a framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology*, *5*(3), 347–358. https://doi.org/10.1037/1076-8998.5.3.347

Grote, G., & Künzler, C. (2000). Diagnosis of safety culture in safety management audits. *Safety Science*, *34*(1–3), 131–150. https://doi.org/10.1016/S0925-7535(00)00010-2

Guldenmund, F. W. (2000). The nature of safety culture: a review of theory and research. *Safety Science*, *34*(1–3), 215–257. www.elsevier.com/locate/ssci

Guldenmund, F. W. (2007). The use of questionnaires in safety culture research - an evaluation. *Safety Science*, *45*(6), 723–743. https://doi.org/10.1016/j.ssci.2007.04.006

Guldenmund, F. W. (2010). *Understanding and Exploring Safety Culture*. BOXPress.

Hamaideh, S. (2004). *Safety Culture Instrument: A Psychometric Evaluation*. Jordan University of Science and Technology.

Health & Safety Executive. (2005). A review of safety culture and safety climate literature for the development of the safety culture inspection toolkit. Research Report 367. *Health (San Francisco)*.

Henriqson, É., Schuler, B., van Winsen, R., & Dekker, S. W. A. (2014). The constitution and effects of safety culture as an object in the discourse of accident prevention: A Foucauldian approach. *Safety Science*, *70*, 465–476. https://doi.org/10.1016/j.ssci.2014.07.004

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Hudson, P. (1999). *Safety Culture – Theory and Practice*. *December*, 1–2.

Hudson, P. (2001a). Aviation Safety Culture. *Safeskies*, 1–23.

Hudson, P. (2001b). Safety management and safety culture the long, hard and winding road. *Occupational Health & Safety Management Systems Proceedings of the First National Conference*, 3. http://www.ohs.com.au/ohsms-publication.pdf#page=11

INSAG. (1991). SAFETY SERIES No . 75-INSAG-4 INTERNATIONAL ATOMIC ENERGY AGENCY , VIENNA , 1991 CATEGORIES IN THE IAEA SAFETY SERIES. In *Safety Series* (Issue 75).

International Civil Aviation Organization. (1993). INVESTIGATION OF HUMAN FACTORS IN ACCIDENTS AND INCIDENTS. *Human Factors Digest #7*.

Johnsen, S. O., Vatn, J., Jersin, E., Veiseth, M., Rosness, R., Lamvik, G., Steiro, T., Hagen, O., & Herrera, I. (2003). Review of existing knowledge applicable to safety culture at interfaces in European Railway undertakings. *SINTEF Report*.

Johnson, S. E. (2007). The predictive validity of safety climate. *Journal of Safety Research*, *38*(5), 511–521. https://doi.org/10.1016/j.jsr.2007.07.001

Kramer, J. (2019). *Deep Democracy*. Jitske Kramer & Boom uitgevers Amsterdam.

Lee, T. (1998). Assessment of safety culture at a nuclear reprocessing plant. *Work and Stress*, *12*(3), 217–237.

Obadia, I. J., Vidal, M. C. R., & e Melo, P. F. F. F. (2007). An adaptive management system for hazardous technology organizations. *Safety Science*, *45*(3), 373–396. https://doi.org/10.1016/j.ssci.2006.07.002

Parker, D. L., Lawrie, M., & Hudson, P. T. W. (2006). A framework for understanding the development of organisational safety culture. *Safety Science*, *44*(6), 551–562. https://doi.org/10.1016/j.ssci.2005.10.004

Parker, D., Lawrie, M., & Hudson, P. (2006). A framework for understanding the development of organisational safety culture. *Safety Science*, *44*(6), 551–562. https://doi.org/10.1016/j.ssci.2005.10.004

Patankar, M. S. (2003). A study of safety culture at an aviation organisation. *International Journal of Applied Aviation Studies*, *3*(1), 243–258.

Patankar, M. S., Brown, J. P., Sabin, E. J., & Bigda-Peyton, T. G. (2012). *Safety Culture; Building and Sustaining a Cultural Change in Aviation and Healthcare*. Ashgate Publishing Limited.

Piers, M., Montijn, C., & Balk, A. (2009). Safety Management System and Safety Culture Working Group (SMS WG) SAFETY CULTURE FRAMEWORK FOR THE ECAST SMS-WG. *ECAST : European Strategic Safety Initiatives*, *March*, 1–14.

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate.

Richter, A., & Koch, C. (2004). Integration, differentiation and ambiguity in safety cultures. *Safety Science*, *42*(8), 703–722. https://doi.org/10.1016/j.ssci.2003.12.003

Rubin, M., Giacomini, A., Allen, R., Turner, R., & Kelly, B. (2020). Identifying safety culture and safety climate variables that predict reported risk-taking among Australian coal miners : An exploratory longitudinal study. *Safety Science*, *123*(December 2019), 104564. https://doi.org/10.1016/j.ssci.2019.104564

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: does CRM training work? *Human Factors*, *43*(4), 641–674.

Schein, E. H. (2004). *Organizational culture and leadership*. Jossey-Bass.

Sexton, J. B., & Klinect, J. R. (2017). The link between safety attitudes and observed performance in flight operations. *Human Error in Aviation*, *1996*, 513–518. https://doi.org/10.4324/9781315092898-25

Silbey, S. S. (2009). Taming Prometheus: Talk about safety and culture. *Annual Review of Sociology*, *35*, 341–369. https://doi.org/10.1146/annurev.soc.34.040507.134707

Sorensen, J. . N. (2002). *Safety culture : a survey of the state-of-the-art*. *76*, 189–204.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2011). *Implementing safety management systems in aviation*. Ashgate.

Taylor, J. C., & Thomas, R. L. (2003). Toward Measuring Safety Culture in Aviation Maintenance: The Structure of Trust and Professionalism. *International Journal of Aviation Psychology*, *13*(4), 321–343. https://doi.org/10.1207/S15327108IJAP1304_01

Thaden, T. L. Von, & Mitchell-Gibbons, A. (2008). The safety culture indicator scale measurement system (SCISMS). *Human Factors Division Institute of Aviation. Technical Report HFD-08-03/FAA-08-2*, *July*, 121.

Vuuren, W. Van. (2000). Cultural influences on risks and risk management: six case studies. *Safety Science*, *34*, 5–9. https://docs.google.com/a/wmu.se/file/d/0B2Ifx_f4ofuXNzlSNlQ2UXV4MGs/edit

Westrum, R. (2004). A typology of organisational cultures. *Quality and Safety in Health Care*, *13*(SUPPL. 2), 22–27. https://doi.org/10.1136/qshc.2003.009522

Wiegmann, D. A., Zhang, H., Thaden, T. Von, Sharma, G., & Mitchell Gibbons, A. (2002). A synthesis

of safety culture and safety climate research. *Aviation Research Lab, Institute of Aviation. Technical Report ARL-02-3/FAA-02-2*, *June*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Modality+Differences+in+Advanced+Cockpit+Displays:+Comparing+Auditory+Vision+and+Redundancy+for+Navigational+Communications+and+Traffic+Awareness#0

Wiegmann, D. A., Zhang, H., Von Thaden, T. L., Sharma, G., & Gibbons, A. M. (2004). Safety culture: An integrative review. *International Journal of Aviation Psychology*, *14*(2), 117–134. https://doi.org/10.1207/s15327108ijap1402_1

Williamson, A. M., Feyer, A., Cairns, D., & Biancotti, D. (1997). THE DEVELOPMENT OF A MEASURE OF SAFETY CLIMATE: THE ROLE OF SAFETY PERCEPTIONS AND ATTITUDES. *Safety Science*, *25*(1), 15–27.

Woods, D. D., Dekker, S. W. A., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Ashgate Publishing Limited.

Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology*, *65*(1), 96–102. https://doi.org/10.1037/0021-9010.65.1.96

Zohar, D. (2000). A group level-model of safety climate: testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology*, *85*(4), 587–596.

Chapter 3. Goal Conflicts, Classical Management and Constructivism: How Operators Get Things Done

Leonie Boskeljon-Horst, Robert Jan de Boer, Simone Sillem, Sidney Dekker (2022)

Safety MDPI 8 (37)

Doi: 10.3390/safety8020037

Goal Conflicts, Classical Management and Constructivism: How Operators Get Things Done

Abstract

In this study we identify the differences in goal realisation when applying two conflicting paradigms regarding rule perception and management. We gathered more than 30 scenarios where goal conflicts were apparent in a military operational unit. We found that operators repetitively utilized certain routines in executing their tasks in an effort to realize several conflicting goals. These routines were not originally intended nor designed into the rules and not explicitly included in documentation. They were not necessarily at odds with the literal wording and / or the intent of rules and regulations although we did find examples of this. Our data showed that local ingenuity was created innovatively within the frame of existing rules *or* kept invisible to those outside the unit. The routines were introduced and passed on informally, and we found no evidence of testing for the introduction of new risks, no migration into the knowledge base of the organisation, and no dissemination as new best practices. An explanation for this phenomenon was found in the fact that the military organisation was applying a top-down classical, rational approach to rules . In contrast, the routines were generated by adopting a constructivist view of rules as dynamic, local, situated constructions of operators as experts. The results of this study suggests that organisations are more effective in solving goal conflicts and creating transparency on local ingenuity if they adopt a constructivist paradigm instead of, or together with, a classical paradigm.

1. Introduction

Rules and regulations are a necessary resource in organisations to help people remember the steps in a task under challenging circumstances, to educate and train people for their jobs, to ensure that people can cooperate effectively, for design and planning purposes, and as a means to identify variances in behaviour that are deemed unacceptable (Hale & Borys, 2013a). However, rules and procedures must not excessively restrict people to cope with the variances that they routinely encounter while they execute tasks (Dekker, 2003a) (De Boer, 2021). Organisations typically suffer from a tension between these two paradigms (Dekker, 2003a) (Hollnagel, 2014) (Mendoza et al., 2020), where management typically considers human behaviour as the cause of incidents and accidents and safety as something to be achieved by designing safe systems and controlling human behaviour within that system (Dekker, 2014b). Employees on the other hand pursue many goals (of which safety is one) by adapting to variable conditions in which they work (Dekker & Tooma, 2021).

1.1 Goal conflicts, violations & local ingenuity

Applicable rules do not accurately represent all of the company's interests such as productivity and efficiency pressures or particular deadlines (Laurence, 2005) (De Boer et al., 2011) (Mendoza et al., 2020). These pressures are often tacit and not sufficiently represented in the rules, but they have an effect on the choices that are made in the moment. People will find shortcuts in tasks to get the job done, often instructed by more experienced colleagues. Because it is impossible to anticipate every possible work situation, it is seen as necessary to infringe on rules to accommodate the varieties in work situations (Woods et al., 2010) (Bye & Aalberg, 2020). In this sense, goal attainment is facilitated by departures from documented rules and procedures. Management condones this, turning a blind eye without alleviating conflicts (Shojania & Dixon-Woods, 2013) (Dekker, 2014b) (Dekker, 2019). Rasmussen (Rasmussen, 1997) has suggested that organisations continuously suffer from pressures of productivity and profitability that undermine safety barriers, and Hollnagel has similarly described trade-offs between thoroughness and efficiency (Hollnagel, 2009).

1.2 Two conflicting views of rules and the covert nature of local ingenuity

Within a top-down classical, rational approach to rules inspired by Reason-thinking ((Reason, 1990) termed model 1 by Hale & Borys (2013a)), rules and procedures are seen as desirable, necessary and unavoidable ways to direct and control human behaviour. Violating these rules and procedures is seen as negative behaviour that needs to be understood in order to be suppressed, as every violation means a first step towards causing an accident (Hale & Borys, 2013a). Compliance to procedures is seen as synonymous for safe behaviour (Dekker, 2018), because procedures control erratic human behaviour (Dekker, 2006). Enhancing knowledge on procedures and ensuring compliance leads to progress on safety (Dekker, 2014b). Hale & Borys (2013a) have listed the strengths and weaknesses of model 1, reproduced with elaboration from the paper in Appendix 1.

An alternative view of rules is constructivist, viewing rules as dynamic, local, situated constructions of operators as experts (model 2 (Hale & Borys, 2013a)). In this paradigm the reason for violating procedures is found in that multiple, conflicting and implicit goals have to be achieved (Dekker, 2014b). Dekker (Dekker, 2018) suggested a complex and unpredictable world of work in which a difference between the way work is assumed or imagined to be done and the way it is actually done is unavoidable. Rules and work are not synonyms and rules are just one resource used by employees (Dekker, 2003a). Under duress of goal conflicts employees are able to execute their work because they are able to innovate and improvise outside the rules and procedures (Dekker, 2018). Particularly, being more efficient or more thorough than expected, being able to deal with a lack of resources and delegating safeguards are ways found to get the work done (Ashour et al., 2021). Within this view on violations, rules and procedures are seen as local behavioural patterns that emerge from experience. Because rules can never completely cover every circumstance, translation and adaptation are inevitable and violations are necessary when rules and reality do not match (Hale & Borys, 2013a). Hale & Borys (2013a) have listed the strengths and weaknesses of model 2, reproduced with clarification from the text in Appendix 1.

Mendoza et al. (Mendoza et al., 2020) found that management and operators had divergent expectations of how procedures are used, when they're most useful, and reasons why operators don't utilize the

procedure amendment process. In essence management tend to adhere to a top-down classical, rational approach to rules (model 1) whereas operators tend to take a constructivist view, seeing themselves as sufficiently experienced to deviate where they feel this is necessary (model 2). Hale and Borys suggest a framework which combines strengths of model 1 and 2 which "places the monitoring and adaption of rules central to its management process and emphasises the need for participation of the intended rule followers in the processes of rulemaking, but more importantly in keeping those rules alive and up to date in the process of regular and explicit dialogue with first-line supervision" (Hale & Borys, 2013b).

1.3 Local ingenuity

As a result of the tension between the two paradigms of rule perception and management, routines exist that solve goal conflicts and become part of operators' regular repertoire but are relatively invisible to management (De Boer, 2021). These routines might be a source of pride but are invisible to all but the inner crowd: management's paradigm of full compliance effectively stifles any understanding and recognition for alternative routines. Rather, there may be repercussions and at least a push-back on the routine, making it more difficult for the operators to achieve high goal realisation in future. We expect these routines to be repeatedly applied, passed on from employee to employee and therefore relatively stable.

In contrast to many previous authors we have dubbed these routines 'local ingenuity' (rather than non-compliances, deviations, gaps or shortcuts), stressing the non-normative perspective that this research requires us to take. Note that in our definition of local ingenuity these are not necessarily at odds with the literal wording, or the intent, of rules and regulations; however the routines were not originally intended, have not been designed into the rules and are not explicitly included in the current documentation. Because of their covert nature, examples of local ingenuity have likely not been assessed for risks, are not absorbed in the knowledge base of the organisation, and have not been disseminated as best practice.

1.4 Study objectives

The overall aim of the present study is to identify how the two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013a) affect goal realisation and the resolution of goal conflicts in an operational setting. We aim to identify examples of (possibly covert) routines to solve goal conflicts (called local ingenuity) and confirm the contribution of these routines to goal realisation and the resolution of goal conflicts. We aim to identify whether these routines are known to management and whether they are aligned with the literal wording or the intent of rules and regulations. We will investigate whether they have been assessed for risks, have been absorbed in the knowledge base of the organisation, or have been disseminated as best practice. We aim to identify the factors that aid or thwart the creation of these routines – in particular the paradigms that are adopted by management and operators about rules.

The main research question that we address is: what are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013a)? The sub research questions are:

- Can examples of goal conflict and of local ingenuity be identified within an operational environment?
- Is the sum of possible goal realisation improved in the examples of local ingenuity compared to the base case?
- Are these examples of local ingenuity known to management?
- Are the examples of local ingenuity aligned with the literal wording or the intent of rules and regulations?
- How is local ingenuity absorbed into the existing knowledge base of the organisation and how is it tested for the introduction of new risks?
- What rule paradigms have been applied in the base case with goal conflicts and in the creation of local ingenuity? How have these paradigms influenced the creation of local ingenuity?

These research questions have led the research initiative within an operational unit of a military organisation.

2. Method

To study how two conflicting paradigms regarding rule perception and management affect goal realisation in an operational setting through the identification of local ingenuity, the authors chose to limit the research to the case study of a single target organisation. We chose a qualitative survey research methodology to enable collection of data over a relatively widespread sample of varying seniority of the employee population, supported by numerical analysis where appropriate.

2.1 Target organisation

The study was conducted in a flight squadron within the Royal Netherlands Air Force (RNLAF) that was sought out because it was felt to be successful in creating local ingenuity. The RNLAF has formulated the ambition to improve its competitive advantage under the umbrella term "5th generation air force" and it is thought that ways to improve goal realisation and resolve goal conflicts contributes to this initiative. Although the military nature of the operation means that not all research data can be made public, the researchers found the target organisation in initial talks to be more open about goal conflicts and possible violations than most commercial organisations.

The flight squadron is an operational unit encompassing 55 employees and around thirty aircraft. The squadron activities include flight planning, briefing and debriefing, line maintenance and ancillary support activities. The squadron's home base is one of the Dutch air fields but it is also active in other NATO countries for joint exercises, and on deployment missions elsewhere. The goals of the squadron under study are to maximise the individual's potential by developing his/her skills and building their expertise, executing the missions assigned to the squadron by the air force commander and his delegates within

the appropriate budget, and doing this safely. An overall RNLAF goal is being compliant with existing regulations, rules and procedures. For this research we have translated these objectives into four organisational goals: safety/security, building expertise (developing skills), productivity (expected performance) and compliance to rules and regulations.

2.2 Interviews

To collect data, semi-structured interviews were held with twelve people. The interview guide included two main questions:

- Could you tell us about something that makes executing your job difficult?
- How do you deal with this?

All interviews were preceded with written and oral explanation of the aims and methods of the study, the way data was anonymised and stored, and how the study results are disseminated. Each participant signed an informed consent form attesting to their understanding and approval of the study. At least two of the three researchers were present during each interview. The interviews were not recorded. Notes were taken, processed and validated by the interviewees for correct reflection. Notes were deleted and only the validated and anonymised interview data is stored.

2.3 Scenario descriptions

The interviews resulted in 'scenario descriptions': a narrative describing a (base) case with goal conflicts, and (if applicable) a solution in which the sum of possible goal realisation was improved. We identified these solutions as examples of local ingenuity if they were aimed at solving goal conflicts, were not originally intended, were not designed into the rules and were not explicitly included in the current documentation and became part of the regular repertoire of the operator. We identified whether these examples were at odds with the literal wording, or the intent, of existing rules and regulations. All scenarios were described using the behavioural and psychological features Hollnagel et al (Hollnagel et al., 2022) proposed as guidance to analyse situations. In case different interviewees came up with the same base case, the stories were combined into one scenario. We excluded examples that were not directly pertaining to the target organisation. All scenarios were stored on a secure network, only accessible by the three researchers.

2.4 Coding

For each scenario it was identified which goals were being pursued in the base case, and how goal realisation changed in case of local ingenuity. A goal conflict was defined when one of the four organisational goals hampers the attainment of at least one of the other three goals. To determine the extent of improved goal attainment as a result of local ingenuity, a 5-point Likert scale was used with a mid-point neutral score. Scale labels for the safety/security goal ranged from severely more risk to severely less risk. For the goals of productivity and expertise the scale labels ranged from greatly

diminished to greatly improved. For each scenario the non-compliance was determined for the base case and the local ingenuity (if available) using a simple yes/no code. Non-compliance was considered to be either a breach of procedure or a violation.

To discern which paradigm regarding rule perception and management was utilised in each scenario for base case and local ingenuity, we used the strengths and weaknesses (characteristics) as described by Hale & Borys (Hale & Borys, 2013a). The application of these to identify the dominant paradigm in an operational setting is (as far as the authors are aware) novel. Clarification for the strengths and weaknesses of model 1 and 2 was derived from the original paper and made available to the researchers (see Appendix 1). For each base case and each local ingenuity separately the researchers scored the applicability of a characteristic using a 3-point Likert scale ranging from "the characteristic was apparent in the narrative" to "clear evidence that the specific characteristic was not applicable". The neutral point was labelled "doubt regarding the applicability of a characteristic". Cronbach's alpha was calculated to determine the internal consistency of the individual characteristics for each model. A high alpha result means the characteristics contribute evenly to the overall applicability (scenarios x characteristics of either model 1 or model 2). An $\alpha \geq 0{,}8$ is usually considered acceptable (Bryman, 2016).

An interrater reliability was calculated for all variables using Fleiss' kappa, as there were more than two coders. Although clear guidance is unavailable on the appropriateness of values for Fleiss kappa (Sim & Wright, 2005), the authors deemed a Kappa $\geq 0{,}5$ as acceptable.

3. Results

3.1 Interviews

Interviews were conducted with twelve employees of the target organisation, nine officers and three non-commissioned officers. All interviewees were males between the age of 25 and 45. An interview took on average one hour. After finalizing the narratives, these were validated by the interviewees, giving them the opportunity to correct any factual mistakes.

3.2 Scenarios

A total of 33 scenarios remained after testing against the criteria (subject to goal conflicts; pertaining to the target organisation); five were eliminated. Of these, 24 were directly related to the main task (flight operations and (training)missions) of the unit and nine were pertaining to maintenance and/or support. We considered the scenarios to constitute a non-compliance with existing rules, regulations or guidelines in three of the base cases (9%, $\kappa = 0.86$). In two cases these non-compliances were not visible to the organisation external to the unit studied. Four example scenarios are reproduced in table 1.

Table 1: Four example scenarios

Scenario 1: *Requesting terrain for military training is accompanied by a quite a bit of administrative work and takes at least two weeks to complete. Following the prescribed procedures hampers the productivity of the unit and significantly reduces flexibility. To cope, all the necessary information to request terrain is summarised in one A4 page which is submitted to the authorities. This local ingenuity is within the existing rules. A copy is taken by the employees on the exercise. The productivity goal is hereby greatly improved. For the base case the model 1 characteristics were predominantly applicable and the model 2 characteristics were not. The characteristics of model 2 are applicable to the local ingenuity. Some of the model 1 characteristics are applicable to the local ingenuity, some are not and some can neither be confirmed nor disproved.*
Scenario 12*: From civil airports it is not allowed to take off after sunset. Civil airports are used as an alternate in case of an emergency. It this is the case, there is a chance that once landed it will not be possible to leave on the same day, creating the challenge of securing the aircraft, disrupting productivity and impacting the main task of the unit. No local ingenuity accompanies this base case. No non-compliance was found. Model 1 characteristics are applicable to the base case and model 2 characteristics are not.*
Scenario 13: *Military personnel may have to use violence to attain a military goal. Rules of engagement (ROE's) describe the circumstances, conditions, degree and manner of violence that is allowed. Interpreting the ROE's can be difficult in some situations, resulting in the risk of unwantedly breaking these rules. The resulting conservative attitude negatively impacts the safety and expertise goals. Within the existing rules a briefing is constructed by an employee, which provides more background and clarity of the ROE's. This briefing results in less safety risk and improved building of expertise. Model 1 characteristics are applicable to the base case and model 2 characteristics are not. The reverse is found for the local ingenuity.*
Scenario 28*: Employees are obligated to have certain rules and procedures physically with them when training, of which most are digitally stored on Ipads. Finding the right procedure during training takes time and effort, negatively impacting the safety and productivity goals. Within the existing rules, a master pdf file is constructed, which encompasses short cuts to all the necessary rules. This turns the digital documents into one master file. Both the safety and productivity goal are improved, diminishing the goal conflict with the compliance goal. No procedures pertain to either the base case or the local ingenuity, resulting in model 1 and 2 characteristics neither confirmed nor disproved for this scenario.*

3.3 Local ingenuity

We found examples of local ingenuity within 22 of the 33 scenarios. These were screened against the criteria (aimed to alleviate to goal conflicts; not originally intended or designed into the rules; not

explicitly included in the current documentation; part of the regular repertoire of the operator; pertaining to the target organisation) and none were eliminated.

We considered the examples of local ingenuity to constitute a non-compliance with existing rules, regulations or guidelines in twelve cases (55%, $\kappa = 0.94$). This varied from the breach of a self-determined learning goal to the violation of a flight rule. In seven of these twelve cases the examples of local ingenuity were not visible to the organisation external to the unit studied. There were no cases where a non-compliance in the base case was changed into a compliancy by the local ingenuity

3.4 Goal conflicts

For all but two scenarios, a goal conflict between the compliance goal and at least one of the other organisational goals (safety, productivity and building of expertise) was determined, impeding the achievement of these latter goals. The interrater reliability for determining which of the organisational goals is negatively impacted by the compliance goal was sufficient ($\kappa \geq .85$ for each organisational goal).

As per definition, goal attainment opportunities were found in each scenario. A total of 54 opportunities were found of which fifteen for the safety/security goal, nine for the building of expertise goal and 30 for the productivity goal. We found that in nearly all scenario's (30 of the 33) productivity was negatively impacted by the compliance goal and in most scenario's (20 of the 33) the attainment of two organisational goals was hampered by the compliance goal. Most common combination found was safety/security and productivity (twelve instances), followed by building expertise and productivity (six instances) and safety/security and building expertise (two instances). In one instance all three goals were negatively impacted by the compliance goal.

The 22 examples of local ingenuity had a positive effect on goal attainment, particularly building expertise (six instances) and productivity (19 instances). The change was either 'improved' or 'greatly improved'. Not in all cases was the riskiness of the local ingenuity increased versus the base case; in fact in some cases it was reduced. In four instances, the local ingenuity improved the attainment of the safety/security goal. In four other instances however, the local ingenuity increased the safety/security risks. In all cases of an increased risk this was limited to 'slight'. Local ingenuity within the bounds of existing rules and regulations resulted in improvement of the safety, expertise and productivity goals in respectively three, five and ten times out of the 20 attainment opportunities. Where local ingenuity constituted a non-compliance with existing rules, regulations or guidelines, it resulted in improvement of the safety, expertise and productivity goals in respectively one, one and eight times out of the 14 attainment opportunities and resulted in an increased safety/security risk for three of the 14 opportunities.

3.5 The characteristics of model 1 and model 2 for the base cases

Our research endeavoured to identify rule management paradigms by using model 1 (top-down classical, rational approach to rules) and model 2 (constructivist view) characteristics. In this section we report the results of our analysis on the base cases.

Based on the coding, we found that the characteristics of model 1 were predominant for most base case descriptions (26 out of 33, κ ≥ 0.56, see table 2). Two of these base cases had no existing rules or procedures that are applicable, hence the model 1 characteristics cannot be used. For the remaining five base cases the model 1 characteristics could neither be confirmed nor disproved. With regard to model 2 we found that the characteristics were *not* applicable for most base case descriptions (27 out of 33, κ ≥ 0.63). For the other six base cases the model 2 characteristics could neither be confirmed nor disproved because no existing rules or procedures were applicable.

Cronbach's alpha was calculated to determine the internal consistency of the characteristics and showed good overall internal consistency for model 1 (α = 0.93) and model 2 (α = 0.99). However, two characteristics of model 1 seemed less relevant to our data set. *Proven effectiveness for simple, 'golden rules' (Behavioural Based Safety)* (κ = 0.27) and *Emphasises the role of organisational complicity in rule violation* (κ = 0.85) were predominantly scored neutral. The low Kappa value of the former is explained by the high agreement between coders in 30 of the 33 cases.

Table 2: results of model 1 and model 2 for the base cases

Model 1														Model 2									
	Strengths							Weaknesses						Strengths						Weaknesses			
scenario	1	2	3	4	5	6	7	8	9	10	11	12	Kappa	13	14	15	16	17	18	19	20	21	Kappa
1	+	0	+	+	+	0	0	0	0	+	+	+	0,88	-	-	-	-	-	-	-	-	-	1
2	+	0	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
3	+	0	+	+	+	0	0	+	0	+	+	+	0,65	-	-	-	-	-	-	-	-	-	1
4	+	0	+	+	+	0	0	+	0	+	0	+	0,52	-	-	-	-	-	-	-	-	-	-0,29
5	0	+	+	+	+	0	0	+	0	+	+	0	0,08	-	-	-	-	-	-	-	-	-	-0,32
6	0	0	0	0	0	0	0	0	0	+	+	0	0,07	-	-	-	-	-	-	-	-	-	-0,50
7	+	+	+	+	+	+	+	+	+	+	+	+	-0,03	-	-	-	-	-	-	-	-	-	1
8	+	+	+	+	+	0	0	+	+	+	+	+	0.82	-	-	-	-	-	-	-	-	-	1
9	+	+	+	+	+	0	0	+	+	+	+	+	0,82	-	-	-	-	-	-	-	-	-	1
10	+	0	+	+	+	0	0	+	+	+	+	+	0,84	-	-	-	-	-	-	-	-	-	1
11	+	-	+	+	+	0	+	+	+	+	+	+	0,83	-	-	-	-	-	-	-	-	-	1
12	+	+	+	+	+	0	0	+	+	+	+	+	0,82	-	-	-	-	-	-	-	-	-	1
13	0	+	+	+	+	0	0	+	+	+	+	+	0,53	-	-	-	-	-	-	-	-	-	1
14	+	+	+	+	+	0	0	+	+	+	+	+	0,82	-	-	-	-	-	-	-	-	-	1
15	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
16	+	0	+	+	+	0	0	+	0	+	+	+	0,88	-	-	-	-	-	-	-	-	-	1
17	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
18	+	0	+	+	+	0	0	+	0	+	+	+	1	-	-	-	-	-	-	-	-	-	1
19	+	0	+	+	+	0	0	+	0	+	+	+	1	-	-	-	-	-	-	-	-	-	1
20	+	-	+	+	+	0	0	+	-	+	+	+	1	-	-	-	-	-	-	-	-	-	-0,08*
21	0	0	0	0	0	0	0	0	0	0	0	+	-0,13	0	0	0	0	0	0	0	0	0	-0,29
22	0	0	0	0	0	0	0	0	0	0	0	0	-0,03*	0	0	0	0	0	0	0	0	0	1
23	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
24	+	0	+	+	+	0	0	+	0	0	+	+	0,31	-	-	-	-	-	-	-	-	-	-0,16
25	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
26	+	0	+	+	+	0	0	+	0	+	+	+	1	-	-	-	-	-	-	-	-	-	1
27	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
28	0	0	0	0	0	0	0	0	0	0	0	0	-0,10	0	0	0	0	0	0	0	0	0	-0,23
29	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
30	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
31	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
32	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
33	+	+	+	+	+	0	0	+	+	+	+	+	1	-	-	-	-	-	-	-	-	-	1
Kappa	0,90	0,70	0,6	0,9	0,8	0	0,9	0,6	0,8	0,7	0,7	0,8		0,7	0,7	0,6	0,8	0,80	0,7	0,8	0,6	0,8	
* The observed and expected agreement resemble one another, resulting in a K<.5 and mispresenting the level of agreement.																							

3.6 The characteristics of model 1 and model 2 for the local ingenuity

In this section we report the results of our analysis on the local ingenuities. We found that the characteristics of model were predominantly *not* applicable for the examples of local ingenuity (17 out of 22, κ ≥ 0.62, see table 3). For the remaining five local ingenuities the model 1 characteristics could neither be confirmed nor disproved, because no existing rules or procedures were found to be applicable. We found that the characteristics of model 2 were predominant for most of the local ingenuity descriptions (18 out of 22, κ ≥ 0.53). For the remaining four local ingenuities the model 2 characteristics could neither be confirmed nor disproved because no applicable rules or procedures were found.

Cronbach's alpha was calculated to determine the internal consistency of the characteristics and showed good overall internal consistency for model 1 (α = 0.86) and model 2 (α = 0.91). Again, the characteristic *Proven effectiveness for simple, 'golden rules' (Behavioural Based Safety)* seemed less relevant to our data set (κ = 0.15, explained by the high agreement between coders).

Table 3: results of model 1 and model 2 for the local ingenuity. Grey cells represent the scenarios with local ingenuities as a result of perceived freedom within existing rules; white cells represent scenarios with local ingenuities that is not visible, the exception being scenario 11.

	Model 1													Model 2									
	Strengths							Weaknesses						Strengths						Weaknesses			
scenario	1	2	3	4	5	6	7	8	9	10	11	12	Kappa	13	14	15	16	17	18	19	20	21	Kappa
1	+	0	-	+	+	0	0	-	0	-	-	-	1	+	+	+	+	0	+	-	+	-	1
2	-	-	-	-	-	0	+	+	-	-	-	-	0,85	+	+	+	+	+	+	+	+	+	-0,03*
3	-	-	-	-	-	0	+	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	-0,03*
4	-	-	-	-	-	0	0	-	-	-	-	-	0,57	+	+	+	+	+	+	+	+	+	-0,09*
7	-	-	-	0	-	0	0	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	1
9	-	-	-	-	-	0	0	-	-	-	-	-	0,69	+	+	+	+	+	+	+	+	+	-0,04*
11	-	-	-	-	-	0	+	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	1
13	-	0	-	-	0	0	0	-	0	-	-	-	0,62	+	+	+	+	+	+	0	-	-	0,34
14	-	-	-	-	-	0	+	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	1
16	-	-	-	-	-	0	0	-	-	-	-	-	0,82	+	+	+	+	+	+	+	+	+	-0,06*
18	-	-	-	-	-	0	+	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	1
19	-	-	-	-	-	0	+	-	-	-	-	-	0,83	+	+	+	+	+	+	+	+	+	1
20	-	-	-	-	-	0	-	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	-0,35
21	0	0	0	0	-	0	0	0	0	0	0	0	0,54	0	0	0	0	0	+	0	0	0	0,71
22	0	0	0	0	-	0	0	0	0	0	0	0	1	0	0	0	0	0	+	0	0	0	1
24	+	0	-	-	+	0	0	-	0	-	-	-	0,25	+	+	+	+	0	+	-	-	-	0,24
25	0	0	-	-	+	0	0	-	0	-	0	-	0,13	+	+	+	+	0	+	-	-	-	0,63
26	-	-	-	-	-	0	+	-	-	-	-	-	1	+	+	+	+	+	+	+	+	+	1
27	0	0	0	0	-	0	0	0	0	0	0	0	1	0	0	0	0	0	+	0	0	0	0,46*
28	0	0	0	0	-	0	0	0	0	0	0	0	0,34	0	0	0	0	0	+	0	0	0	0,40
29	-	-	-	-	0	0	-	-	-	-	-	-	0,77	+	+	+	+	+	+	+	+	+	1
30	-	-	-	-	-	0	+	-	-	-	-	-	0,61	+	+	+	+	+	+	+	+	+	1
Kappa	0,7	0,8	0,90	0,8	0,5	0,15'	0,9	0,8	0,6	0,7	0,7	0,90		0,90	0,90	0,8	0,80	0,6	-0,03	0,7	0,7	0,5	
* The observed and expected agreement resemble one another, resulting in a Iκ <.5 and misrepresenting the level of agreement.																							

4. Discussion

In this study we aimed to identify what the differences were with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys

(2013a). We report on the findings of a study at an operational squadron of a military organisation. The results are based on 33 scenario's that have been collected through half-structured interviews, and that have been analysed for goal conflict, goal attainment, non-compliances and the characteristics of model 1 or model 2 thinking (Hale & Borys, 2013a). We show that goal conflicts were frequently solved locally with routines that optimize across multiple conflicting goals, dubbed 'local ingenuity'. These routines are not necessarily at odds with the literal wording, or the intent, of rules and regulations; however the routine is not originally intended, is not designed into the rules and is not explicitly included in the current documentation.

4.1 Scenarios

The 33 scenarios describe a goal conflict between four organisational goals: to improve safety or security, build the expertise of the employees, and/or increase productivity or budget utilisation. These base cases were seen as situations where (according to those involved) there was a potential for improvement of the realization of these goals. Local ingenuity was used to realise those goal improvement opportunities for the flight, (training)mission and maintenance/support base cases as expected (Bye & Aalberg, 2020) (Hollnagel, 2018) (Woods et al., 2010).

4.2 Local ingenuity

In 22 of the 33 scenarios that were collected an adjustment to work was identified to alleviate tension between conflicting goals (i.e. local ingenuity). These examples support Hollnagel (Hollnagel, 2014) and Dekker (Dekker, 2018) who described gaps between what is written about how work should be done and what is actually done, and how this is a result of the efforts to capture a non-deterministic and complex world into rules and procedures. In some cases (45%) we found that employees were able to solve goal conflicts within existing rules and procedures. However in 55% of cases (12 out of 22) some rule, regulation or guideline was violated; in many cases (7 out of 12) because the solution was not visible outside the target organisation. These results suggest that the process of "keeping the rules alive and up to date in the process of regular and explicit dialogue with first-line supervision" (cf. Hale & Borys (2013b)) is failing. Irrespective of whether local ingenuity was compliant with existing rules and regulations or a non-compliance, we saw limited proof of it being tested for (new) risks, it being incorporated into rules and regulations or it being disseminated to other units. Exceptions were scenario 11, of which the local ingenuity was tested for (new) risks and scenario 13, of which the local ingenuity was shared with other units.

4.3 Goal conflicts and goal opportunities

We found goal conflicts between the compliance goal and the three other organisational goals in all of the scenarios as predicted by the literature (Rasmussen, 1997). Applying local ingenuity in the 22 of 33 scenarios led to an overall increase in goal attainment of safety/security, productivity and/or building expertise, regardless whether the local ingenuity was based on perceived freedom within rules or on

the lack of visibility. Not in all cases was the riskiness of the local ingenuity increased versus the base case; in fact in some cases it was reduced.

To illustrate the goal conflict between the internal compliance goal and the organisational goals as well as the effect of local ingenuity on goal attainment, we plotted both the base cases and the local ingenuities found in two 2x2 frameworks (De Boer, 2016), see figure 1 and figure 2.

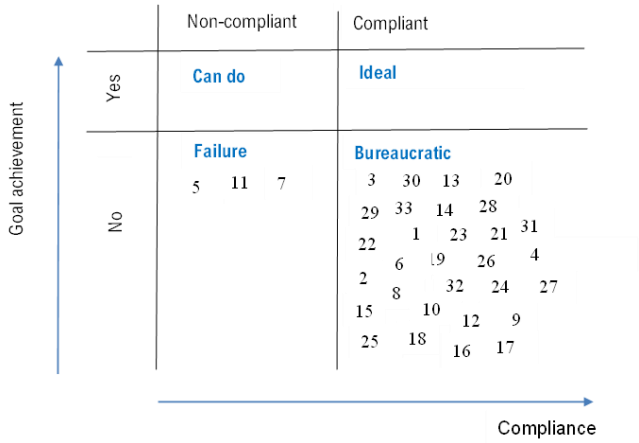

Figure 1. Compliance x goal achievement for the base cases (N=33)

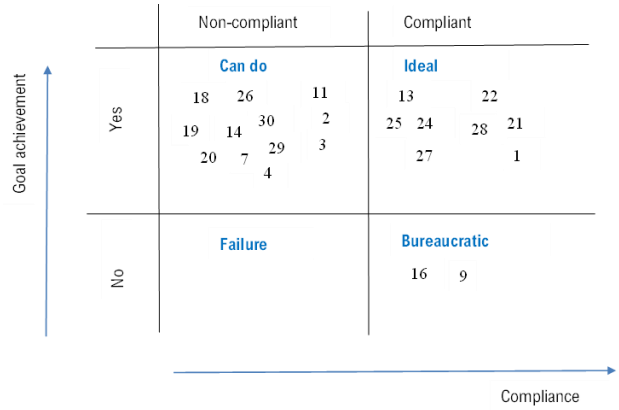

Figure 2. Compliance x goal achievement for local ingenuity (N=22)

As figure 1 shows, the conflict between the compliance goal and the other organisational goals resulted in compliance in 30 out of the 33 three scenarios, but goals were not achieved in any of these cases. The local ingenuity of the employees resulted in 20 out of the 22 instances in goal attainment (figure 2).

The resulting non-compliance was the consequence of the desire to seize goal improvement opportunities. Being compliant and achieving goals, however, is not necessarily an ideal situation: in some cases the intent of the rules was not complied with although in a literal sense there was compliancy. In other cases we found that the rule base was not challenged or modified despite examples of local ingenuity that should have triggered a rules change (local ingenuities 1, 21, 22 and 28).

In only three base cases we encountered non-compliances, whereas we found twelve as a result of the local ingenuities. The quadrupling of the violations was expected, as they were the result of a quicker way to do the job, the design made compliance difficult or even impossible (Lawton, 1998), rules lacked an understanding of the working situation (Laurence, 2005), conflicting demands (Battmann & Klumb, 1993) (Weichbrodt, 2015), and a lack of worker involvement in rulemaking (Kanse et al., 2018). In those cases that local ingenuity led to non-compliances, we identified a small increase in the safety risk with three violations. However, we also identified a decrease in the safety risk with two other non-compliances and an increased improvement in productivity in all non-compliances.

4.4 Model 1 and model 2

According to the interviewees, the lack of transparency on local ingenuity in this organisation was heavily compelled by an approach to rules and compliance that aligns with the characteristics of model 1 (Hale & Borys, 2013a). We found that in the organisation's approach to potential for improvement of the attainment of organisational goals (i.e. the base case) a model 1 alignment was predominant and there was no alignment with model 2 characteristics. The organisation's approach to local ingenuity met most of the characteristics of model 2, and few of the characteristics of model 1. Goal conflicts are more likely and more persevering with model 1 as a result of the rigid way of looking at work. Only by taking on a model 2 perspective was local ingenuity stimulated, and therefore the achievement of conflicting goals possible. In those cases where a model 2 approach was not possible (e.g. due to high visibility outside the unit studied), local ingenuity was smothered and goal achievement seemed sub-optimal. In those cases where local ingenuity was evident however, we found that there was a lack of model 1 thinking, resulting in a lack of clear procedures to support task execution and to train others, and for organisational design purposes (Hale & Borys, 2013a). In case of non-compliances as a result of the local ingenuity, these were continuous and went entirely unmanaged except at the unit level, effectively leading to complicity of superiors.

The characteristic *Proven effectiveness for simple, 'golden rules' (Behavioural Based Safety)* of model 1 turned out to be problematic when coding the scenarios in the target organisation. We found this characteristic in only one base cases and none of the local ingenuities. Behavioural Based Safety is a concept not embraced within the RNLAF, explaining this result. The characteristic *Emphasises the role of organisational complicity in rule violation* of model 1 also seemed to pose a problem. The lack of applicability with the base cases is explained by a lack of non-compliances. For local ingenuity, this characteristic is found to be applicable in six cases of non-compliances.

4.5 Effectiveness of a workplace

The results of our study confirm downward pressure on local ingenuity as a consequence of a predominant focus on model 1 thinking to the exclusion of model 2. The opportunity for local ingenuity is limited to caveats in the rules and regulations or the local ingenuity is invisible to others outside the operational unit (including the rest of the military organisation). Because local ingenuity (be it compliant or not) is largely invisible, it is not tested for (new) risks, not incorporated into rules and regulations or the Safety Management System and not disseminated to other units.

In some instances we found merely cosmetic compliance (i.e. to the literal wording of rules, not the intention). When compliance becomes impossible while trying to solve problems, employees come up with entirely different task executions, that provide a way to work around the problems and still fall within the rules, just different rules than the problems conflict with.

4.6 Theoretical contribution

Our research contributes to the existing literature on model 1 and model 2 thinking by showing that the characteristics can be used to discern which of the paradigms about rules have been adopted in a particular situation. This novel application has empirically confirmed the existence of two conflicting paradigms for rule management, a top-down classical, rational approach to rules (model 1 according to Hale & Borys (2013a)) and a constructivist view (model 2).

We found that one needs to tailor the taxonomy of goals to an organisation when conducting an investigation into goal conflicts, rather than subsuming to the generic classifications in the literature. Our taxonomy of safety/security, productivity and building expertise turned out to be useful to illustrate goal conflicts in all 33 scenarios. Adding the overall internal goal of being compliant to our taxonomy seemed more illustrative than the simple bi-polar ETTO assumption (Hollnagel, 2009) and more tailored to this organisation than the Dynamic Safety Model (Rasmussen, 1997), which encompasses only organisational goals.

4.7 Practical contribution

Our research contributes to practice by adding more nuance when comparing compliance to non-compliance. The findings show that a predominant focus on model 1 to the exclusion of model 2 hampers the effectiveness of the unit studied. Being compliant and achieving goals is not necessarily an ideal situation: in some cases the intent of the rules was not complied with although in a literal sense there was compliancy. In other cases we found that the rule base was not challenged or modified despite examples of local ingenuity that should have triggered a rules change. With this research a way to identify and describe the effectiveness of the entire RNLAF becomes possible. It also provides a way to enhance the effectiveness of the organisation.

4.8 Recommendations

Based on our findings, the military organisation is advised to consider an approach to rules and regulations and local ingenuity that combines the strengths of model 1 and model 2. This includes well-defined autonomy (freedom-in-a-frame), a responsive attitude to rule modifications, operator input, and (dynamic) management of exceptions to meet the complexity of the operation.

Hidden local ingenuity prevents a discussion regarding its appropriateness. Compliant but constrained local ingenuity cannot be deliberately applied by an organisation. Both situations are undesirable from an organisational perspective. The examples of local ingenuity – and the eleven cases in which no local ingenuity was devised despite a goal conflict – justify that rules should allow for a certain autonomy (freedom-in-a-frame) such as proposed by De Boer (De Boer, 2021), and that a balance needs to be found with the five goals of rules (Hale & Borys, 2013a). In our study, the required equilibrium had not been found as in none of the cases the local ingenuity had been integrated into the rule base. In fact, this equilibrium has not been actively sought, parties instead opting in nine cases to keep the local ingenuity within the unit and therefore hidden from entities outside the own unit.

As a result of our study we hope to be able to close the gap between local ingenuity and the rules and regulations, by testing the local ingenuity for appropriateness under specific conditions and integrating it into the rule base (De Boer, 2021). The steps and success of this path will be reported in a next paper.

4.9 Next steps

Hale and Borys (2013a) identified characteristics of both model 1 and model 2, and indicated whether they considered these strengths or weaknesses. The authors suggest a framework which combines strengths of model 1 and 2 which "places the monitoring and adaption of rules central to its management process and emphasises the need for participation of the intended rule followers in the processes of rulemaking, but more importantly in keeping those rules alive and up to date in the process of regular and explicit dialogue with first-line supervision…" (Hale & Borys, 2013b). The authors propose a field test of their approach. De Boer (De Boer, 2021), building on the work of Hale and Borys, detailed how such a framework might be implemented. The author also indicated how significant changes to working practices might first be tested under controlled circumstances to check for unwanted side effects (De Boer, 2021).

5. Conclusions

In this study we aimed to identify what the differences were with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013a). We gathered more than 30 scenarios where goal conflicts were in evidence within an operational unit of a military organisation.

Goal conflicts were frequently solved locally with routines that optimize across multiple conflicting goals, dubbed 'local ingenuity'. These routines were not necessarily at odds with the literal wording, or the

intent, of rules and regulations; however the routine was not originally intended, was not designed into the rules and was not explicitly included in the current documentation.

In this organisation, local ingenuity was either created within the existing rule base or invisible to those outside the unit, suggesting that the process of "keeping the rules alive and up to date in the process of regular and explicit dialogue with first-line supervision" is failing. The examples of local ingenuity were not tested for the introduction of new risks, did not become part of the knowledge base of the organisation, and were not disseminated as best practices.

An explanation for this phenomenon was found in the fact that the military organisation, using the wordings of Hale and Borys (2013a), was applying a top-down classical, rational approach to rules. We found clear evidence that the RNLAF is using the model 1 paradigm in each of the 33 scenarios, to the exclusion of the model 2 paradigm. The model 1 paradigm focuses on a priori devised rules and procedures that encompass the best way to do the work and is unable to encapsulate the dynamical complexity of a work floor. It is expected that the target entity is typical for many other organisations.

In contrast, the examples of local ingenuity followed from a constructivist view of rules as dynamic, local, situated constructions of operators as experts. This shows that the top-down classical, rational approach to rules is insufficient to understand how work is done. The model 1 focus is constraining the development and transparency of local ingenuity and results in the inhibition of goal conflict solutions. The local ingenuity is suboptimal, leading to overall suboptimal effectiveness of the unit studied.

The results of this study suggests that organisations are more effective in solving goal conflicts and creating transparency on local ingenuity if they adopt a constructivist paradigm combined with a classical paradigm. This includes well-defined autonomy (freedom-in-a-frame), a responsive attitude to rule modifications, operator input, and (dynamic) management of exceptions to meet the complexity of the operation.

References

Ashour, A., Ashcroft, D. M., & Phipps, D. L. (2021). Mind the gap: Examining work-as-imagined and work-as-done when dispensing medication in the community pharmacy setting. *Applied Ergonomics*, *93*, 103372. https://doi.org/10.1016/j.apergo.2021.103372

Battmann, W., & Klumb, P. (1993). Behavioural economics and compliance with safety regulations. *Safety Science*, *16*(1), 35–46. https://doi.org/10.1016/0925-7535(93)90005-X

Bryman, A. (2016). *Social research methods*. Oxford University Press.

Bye, R. J., & Aalberg, A. L. (2020). Why do they violate the procedures? – An exploratory study within the maritime transportation industry. *Safety Science*, *123*(November 2019), 104538. https://doi.org/10.1016/j.ssci.2019.104538

De Boer, R. J. (2016). Research into control loop flaws at a maintenance squadron using narratives. *32nd EAAP Conference*.

De Boer, R. J. (2021). *Safety Leadership. A different, doable and directed approach to operational improvements.* CRC Press.

De Boer, R. J., Koncak, B., Habekotté, R., & Hilten, G. Van. (2011). Introduction of ramp-LOSA at KLM Ground Services. *Amsterdam University of Applied Sciences*.

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

Dekker, S. W. A. (2006). *The Field Guide to Understanding Human Error*. Ashgate Publishing Limited.

Dekker, S. W. A. (2014). *Safety Differently. Human factors for a new era.* Apple Academic Press Inc.

Dekker, S. W. A. (2018). *The Safety Anarchist*. Routledge Taylor & Francis Group.

Dekker, S. W. A. (2019). *Foundations of safety science*. Taylor & Francis Group, LLC.

Dekker, S. W. A., & Tooma, M. (2021). A capacity index to replace flawed incident-based metrics for worker safety. *International Labour Review*. https://doi.org/10.1111/ilr.12210

Hale, A., & Borys, D. (2013a). Working to rule, or working safely? Part 1: A state of the art review. *Safety Science*, *55*, 207–221. https://doi.org/10.1016/j.ssci.2012.05.011

Hale, A., & Borys, D. (2013b). Working to rule or working safely? Part 2: The management of safety rules and procedures. *Safety Science*, *55*, 222–231. https://doi.org/10.1016/j.ssci.2012.05.013

Hollnagel, E. (2009). *The ETTO principle*. Taylor & Francis Ltd.

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Hollnagel, E. (2018). *Safety-II in practice Developing the resilience potentials*. Routledge Taylor & Francis Group.

Hollnagel, E., Laursen, T., & Sørensen, R. (2022). A day when (Almost) nothing happened. *Safety Science*, *147*(December 2021). https://doi.org/10.1016/j.ssci.2021.105631

Kanse, L., Parkes, K., Hodkiewicz, M., Hu, X., & Griffin, M. (2018). Are you sure you want me to follow this? A study of procedure management, user perceptions and compliance behaviour. *Safety Science*, *101*(August 2017). https://doi.org/10.1016/j.ssci.2017.08.003

Laurence, D. (2005). Safety rules and regulations on mine sites - The problem and a solution. *Journal of Safety Research*, *36*(1), 39–50. https://doi.org/10.1016/j.jsr.2004.11.004

Lawton, R. (1998). Not working to rule: Understanding procedural violations at work. *Safety Science*, *28*(2), 77–95. https://doi.org/10.1016/S0925-7535(97)00073-8

Mendoza, A. M., Liu, S. C., Dumlao, S. V., Hendricks, J. W., Son, C., Sasangohar, F., & Peres, S. C. (2020). Where Two Ends Meet: Operator and Stakeholder Perceptions of Procedures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 1350–1354. https://doi.org/10.1177/1071181320641322

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science*, *27*(2), 183–213. https://doi.org/10.16250/j.32.1374.2016270

Reason, J. (1990). *Human Error*. Cambridge University Press.

Shojania, K. G., & Dixon-Woods, M. (2013). Bad apples: Time to redefine as a type of systems problem? *BMJ Quality and Safety*, *22*(7), 528–531. https://doi.org/10.1136/bmjqs-2013-002138

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257–268. https://doi.org/10.1093/ptj/85.3.257

Weichbrodt, J. (2015). Safety rules as instruments for organizational control, coordination and knowledge: Implications for rules management. *Safety Science*, *80*, 221–232. https://doi.org/10.1016/j.ssci.2015.07.031

Woods, D. D., Dekker, S. W. A., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Ashgate Publishing Limited.

Appendix 1

Strengths and Weaknesses of model 1 and model 2, based on Hale & Borys [1] accompanied by the interpretation used in this research

	Model 1 Strengths	
1	Emphasis on written nature of rules to facilitate checking by regulators for non-compliances	"Emphasis on the scrutiny of writte documentation"
2	Consequences of violations are clear and explicit	"Emphasis on the detection of non-compliance."
3	Management and SME's are most competent in making the rules, which are imposed on the operators	"Rules are devised by experts that are not part of the workforce, to prevent errors and mistakes of the workforce, who are seen as more limited than the experts in their competence and experience."
4	Based on Scientific Management, a rationalist and prescriptive approach to rules	"Rules are seen as a best practice, the one best way to perform a job."
5	Action rules are used to describe the best way to do the work, which are easy to follow for novices	"Because novices do not yet have their own action rules and/or are not able yet to derive them from process rules, they need, at least temporarily rules from experts imposed on them."
6	Rules are the 'gold standard' for correct behaviour (Behavioural Based Safety)	"BBS, visible in observable rules, schedules of observation and feedback provided, is used to ensure compliance."
7	Organisational complicity as a result of violations in case rule compliance conflicts with productivity demands	"This characteristic is only visible when a non-compliance has been detected."
	Model 1 Weaknesses	
8	Sees operators as dumb robots that have to follow the rules that are imposed on them	"Operators need not to bother themselves with thinking about rules and exceptions, just following them or asking management what to do in case a rule cannot be complied with."
9	Violations and errors are voluntary, negative actions, committed by employees that think they know better	"Violation is always wrong and blame worthy"
10	Sees rule-making as a one-off, static process; rules only need modification when the work changes significantly or an accident happens	"Rules are carved in stone, the threshold for rule modification is rather high."
11	Encountered exceptions can only be dealt with by rule book growth	"If operators cannot comply with a rule, management will come up with new rules as a solution"
12	The focus on written rules results in a gap with reality; if rules do not match reality a modification of reality is needed	"Model 1 starts from the written rules and deduces from them that actions are compliances or violations. Modify reality to match the rules."
	Model 2 Strengths	
1	Operators as seen as the real experts central to rule making	"Operators know the complexitiy of the work and are hence the experts needed for rule making."
2	Key to rule use are experience, expertise, construction of meaning and sense-making	"Using of rules is based on both individual perceptions and judgements as well as group processes in which the applicability of rules is determined."
3	Rule-making is a continuous, dynamic process that is never complete	"Rules need to be adapted and translated to be used."
4	Relates abstract, generic, written rules to the flexible, local interpretation visible in routines	"Tacit rules or emerging expert rules are based on experience and socially constructed. They constitute organisational memory on how to deal with situations and change as a result of new experiences and learning."
5	Recognises the importance of exceptions of written rules and sees adaptations not as violations	Reality is much more diverse than rules give credit for which makes violations inevitable, a positive necessity.
6	Placing centraly experience, competence and ability to adapt	"Experience, competence and ability to adapt guide if, when and how to deviate from procedures."
	Model 2 Weaknesses	
7	The making and changing of rules happens locally which lacks transparancy for regulators and creates learning difficulties for novices	"Rules are not always written, making it difficult to check for risks or for novices to get familiar with them. There is no rule modification process."
8	Undervalues the need for the organisation to explicitly manage the use and development of rules	"Changing the rules and their definition fundamentally to match reality. If rules are not described and/or the way to do something is not transferable, this characteristic is not applicable."
9	Differences in interpretation and competence/experience are not visible	"It is unclear whether "the knowledge to vary and adapt the procedures is [...] present", or whether "mistrust of rules leads to needless violations of them". If the difference between erronous deviation and professional deviation is clear and visible, this characteristic is not applicable."

Chapter 4. Aircrews, rules and the bogeyman: Mapping the benefits and fears of noncompliance

Leonie Boskeljon-Horst, Robert Jan de Boer, Vincent Steinmetz, Sidney Dekker (2023)

Safety MDPI 9 (1)

Doi: 10.3390/safety9010015

Aircrews, rules and the bogeyman: Mapping the benefits and fears of noncompliance

Abstract

Although rules support people while executing tasks, they are not the same as work-as-done. It can be impossible to follow the rules and get the job done at the same time. In this study the objective is to better understand the stakes and interests that lie behind retaining gaps between work-as-prescribed and work-as-done mapping the benefits and fears of noncompliance. The study was conducted along the vertical hierarchy of an operational flight squadron of the Royal Netherlands Air Force. We applied a qualitative survey research methodology using semi-structured interviews, complemented by an investigation of relevant documents. We found a public and political commitment to compliance made by the Dutch Department of Defence which reinforces a cycle of issuing promises followed by pressure to keep the promise. This contradicts the found need for adaptation and freedom to use expertise. The official safety narrative seems to convey a hidden message – bad things happen to bad people, reminiscent of a bogeyman. One opportunity to resolve the situation is a doctrine change, changing prescriptive rules to guidelines.

1. Introduction

Rules and procedures help people to remember steps in a task under varying circumstances, ensure that people can work together effectively, function as organisational memory to indicate how processes work and as a means of identifying deviations that might be problematic (Hale & Borys, 2013a). Some rules and procedures are based on previous experiences which operators have not encountered themselves. But rules and procedures are not the same as work-as-done (Hollnagel, 2012). Work takes place in a context of limited resources and multiple goals and pressures as well as surprises, constraints and novelty. In many cases, it can be impossible to follow the rules and get the job done at the same time (Bieder & Bourrier, 2013) (Dekker, 2003a). Work as actually done is often more complex than that which can ever be captured in rules and procedures (Damoiseaux-Volman et al., 2021a) (Hutchins, 1995). Referring to this as 'workarounds' (B. Morrison, 2015) or 'violations' (Reason, 1990) is judgemental and misguided. Realising that there is a push-pull between centralising (through procedural compliance and top-down control) and decentralising is not new (cf. (Rochlin et al., 1987)). Leaning toward decentralisation, for instance, 19th century Prussian military commanders were urged to not order more supplies than absolutely necessary, and not plan beyond the situations they could foresee, thus empowering their subordinates to deviate and adapt, as long as it helped meet the organisation's or mission's intent (Woods & Shattuck, 2000).

In modern military organisations, particularly those held to the rule of law, media scrutiny and democratic parliamentary accountability, the push-pull can be intense between upholding an image of strict, law-abiding compliance while retaining flexible, adaptive capacity. The conflict typically gets

pushed down into local operating units on the various operational frontlines, with stakeholders up and down the hierarchy silently hoping nothing will happen to blow the conflict into full view (and if it does, blaming individual operator error or violations is still a convenient way out (The Hague: Onderzoeksraad voor de Veiligheid, 2018) (Reason, 2008)). In recent work, we described the necessary adaptations to work-as-prescribed – the description and formalisation of work in rules and procedures (Moppett & Shorrock, 2018) – as 'local ingenuity'. These can congeal into routines that are part of the regular (though 'hidden curriculum') repertoire of operators. They remain relatively invisible to management, were not originally intended and are not included in any existing rule base (Boskeljon-Horst et al., 2022) even though they are part of the informal work system (including informal instruction to novices) (see (Antonsen et al., 2012); (McDonald et al., 2002)). Local ingenuity is an expression of care, commitment, expertise and experience, as well as professionalism and judgement. All are necessary to accomplish a task when rules and procedures are insufficient or actively hostile to achieving goals and completing the mission (Lawton, 1998) (Halbesleben et al., 2008) (Tucker et al., 2020).

Picking up where our recent work on local ingenuity (Boskeljon-Horst et al., 2022) left off, in this present paper we dig more deeply into the interests that lie behind retaining the status quo. There is a gap between work-as-prescribed and work-as-done – the way work is actually done (Moppett & Shorrock, 2018). In our approach, the objective is to better understand (if not eventually offer stakeholders means with which to breach) the stalemate by mapping the benefits and fears of noncompliance and the stakes and interests that lie behind both such noncompliance and its denial.

1.1 Conflicting paradigms

Two paradigms regarding rules and compliance can be distinguished. In the top-down classical, rational approach to rules, rules and procedures are seen as desirable, necessary and unavoidable ways to direct and control human behaviour. Violating these rules and procedures is seen as negative behaviour that needs to be understood in order to be suppressed, as every violation means a first step towards causing an accident (Hale & Borys, 2013a). An alternative approach is constructivist, viewing rules as dynamic, local, situated constructions of operators as experts (Hale & Borys, 2013a). In this paradigm the reason for violating procedures is found in that multiple, conflicting and implicit goals must be reached (Dekker, 2014b). A difference between the way work is assumed or imagined to be done and the way it is actually done is unavoidable (Dekker, 2018). Translation and adaptation are inevitable, and violations are necessary when rules and reality do not match (Hale & Borys, 2013a). Furthermore, rules and procedures can lead to diminishing expertise and even to the situation where people stop thinking all together (Klein, 2011).Under the duress of conflicting goals, employees can do their work because they can innovate and improvise outside the rules and procedures (Dekker, 2018). We conducted this study from the constructivist paradigm regarding rules and compliance, because research shows that the classical paradigm does not solve the gaps between work as done and prescribed. Rules and procedures can never fully describe all possible situations employees will encounter nor fully describe exactly how the work should be done (Besnard & Hollnagel, 2014). Expecting people to adhere to rules that do not match practice does not add to safety; adaptation, even if this means non-adherence, can add to safety (Besnard & Greathead, 2003). The classical paradigm therefore seems to result in a stalemate. Since less is known about the constructivist paradigm it seems worthwhile to study whether or not it resolves the gaps between work as done and prescribed.

1.2 Population of interest

The study was conducted along the vertical hierarchy, starting in a flight squadron and ending with the commander of the Dutch air force, the RNLAF. Within a hierarchal organisation such as the RNLAF what is deemed acceptable and unacceptable (what is needed and allowed) is conveyed through the chain of command. On studying the enduring difference between the way work is done and prescribed, it makes sense to expand the original study (Boskeljon-Horst et al., 2022) to include the entire chain of command.

The flight squadron was originally sought out to study how two conflicting paradigms regarding rule perception and management affect goal realisation in an operational setting through the identification of local ingenuity (Boskeljon-Horst et al., 2022). To understand why the difference between the way work is done and prescribed or is intended to remain as a status quo despite the undesirability, we focused on the same squadron. This required us to include the vertical hierarchy of the RNLAF in our scope, up to and including the commander. We also included the Military Aviation Authority (MAA) in our study, as many rules that govern this squadron's operational activities derive from there. The expansion resulted in three focus groups: operators, senior management (the chain of command) and rule maker/enforcers (the MAA).

The flight squadron is an operational unit containing 55 male employees, aged between 22 and 45, and some 30 aircraft. The squadron activities include flight planning, briefing, debriefing, line maintenance, and ancillary support activities. The squadron's home base is one of the Dutch airfields, but it is also active in other North Atlantic Treaty Organization (NATO) countries for joint exercises, and on deployment missions elsewhere. The goals of the squadron under study are to maximise the individual's potential by developing their skills and building expertise, executing the missions assigned to the squadron by the air force commander and their delegates within the appropriate budget, and performing this safely. An overall RNLAF goal is being compliant with existing regulations, rules, and procedures.

From strict to less strict, the RNLAF has to deal with laws, such as the Aviation Act, and with civil regulations from the International Civil Aviation Organisation, the Joint Aviation Authorities and the European Aviation Safety Agency. Furthermore, the RNLAF is impacted by rules from the European Defence Agency and NATO standardisation agreements (STANAG) and rules from the MAA such as the military aviation regulations, acceptable means of compliance and guidance materials. Finally, there are RNLAF rules regarding the operation of the aircraft, flight operations as well as the exceptions or temporary rules, standard operating procedures, guidelines and checklists.

The RNLAF is part of the Dutch Department of Defence (DoD). The DoD is a political instrument that has suffered from a decade of budget cuts, while political ambition and expected output have remained the same. This has created tension within the DoD: budget cuts led to downsizing – which subsequently led to reduced inflow and many vacancies – and an increasing pressure to deliver the expected output with a reduced work force.

Since 2018 all DoD incidents and accidents that are classified as 'big' in terms of consequences (damage to people, environment and/or equipment) are investigated by the Defence Safety Inspectorate. These investigation reports, as well as the formal response of the DoD to the House of Representatives, are

made public. The recommendations made by the Inspectorate are translated by Central Staff[2] of the DoD into rules, policy and safety measures for the armed forces investigated. Since its establishment, the Inspectorate has conducted five investigations regarding RNLAF incidents and accidents.

The RNLAF has the ambition to improve its competitive advantage under the umbrella term "fifth-generation air force" (5GAF). This means the RNLAF embraces terms such as ownership, continuous learning and reflection, adaptability, resilience, agility and flexibility. In a time when the opponent has access to the same technology and equipment, the RNLAF considers the air force employee key in making the difference (Defensie, n.d.).

2. Methodology

We applied a qualitative survey research methodology complemented by an investigation of relevant documents. We conducted qualitative thematic analysis and triangulation based on the gathered interview and document data, focussed on exnovation (Hollnagel, 2014) to ensure accurate description of the way operators perform day-to-day work. Because only the first author (LB), an inside researcher, conducted the interviews, specific attention was given to the concept of 'researcher-as-instrument' (Barrett, 2007). As the co-authors were external researchers and would also be involved in data analysis, it made sense to first consider the impact of the concept on categorisation, analysis and interpretation of the data (Yoon & Uliassi, 2022). Beforehand, the authors agreed on the interview questions and decided that the interviews would be transcribed to exclude subjective representation. Previous studies show that a researcher must to be both reflective and reflexive to build trust with the interviewees and ensure they provide detailed information (Soh et al., 2020). A trusting relationship was ensured as LB was familiar with the population of interest and knew all the interviewees. LB has worked for the RNLAF for over 23 years, including 17 years as an aviation psychologist. RNLAF operators are used to openly discussing all sorts of topics, including non-adherence to rules, with aviation psychologists.

2.1 Interviews

To collect data, we held semi-structured interviews with twelve people, ranging from pilots of the squadron to the RNLAF commander. All interviewees have a prominent role in the fulfilment of airpower, although the upper ranks have not been active operators for at least 10–15 years. All interviewees are male. The age range is 30-44 years in for the operational unit and 45-59 years among the remaining eight interviewees.

Four of the 12 interviewees were identified by their involvement in the previous investigation. These four represent the ranks that can be obtained in the operational unit. We identified the consecutive commanding officers up to the RNLAF commander as eight additional interviewees for the four primary interviewee. All interviewees were specifically chosen because of their knowledge of enduring

[2] Central Staff is the upper layer of the Dutch Department of Defence, responsible for policy. It directs the ministry's activities, allocates the defence budget and monitors defence spending.

differences between the way work is done and prescribed and/or ability to influence these differences. The interviewees are divided into three groups, rule users (interviewees 1 till 4), leaders (interviewees 5 till 9) and rule makers/enforcers (interviewees 10 till 12). To ensure anonymity all participants are referred to by code (i.e., interviewee 1 etc.).

We approached the interviewees by email, explaining the study objective and the relevance to the organisation. Participation was voluntary and the participants signed an informed consent form attesting to their understanding and approval of the study. The semi-structured interviews lasted approximately one hour and were recorded and transcribed to provide accurate records for analysis. Recordings were deleted and only the anonymised transcriptions were stored.

The interview guide included four main questions as discussion topics:

- Do the interviewees recognise a difference between the way work is done and the way work is prescribed in their own work experience?
- Is a difference between work-as-done and work-as-prescribed considered undesirable?
- Why is the difference between the way work is done and prescribed not solved?
- Can the current set of rules and procedures match the dynamic complexity of flight operations?

Examples of follow up questions are:

- Can you give an example of a difference between the way work is done and prescribed?
- Is a 100% compliance attainable considering the goals that need to be attained?
- Are operators allowed to adapt rules to a situation they encounter?
- Why are operators reluctant to use their discretionary space?
- How can rules and procedures be aligned with the dynamic complexity of flight operations?

During the interview, we presented three examples of goal conflicts between compliance and executing the flight mission and asked interviewees for their opinion and advice. Taken from our past research, the three examples are:

1. When fire fighter capacity is low, airport regulations state that large passenger-carrying helicopters are not allowed to take off or land at the airbase. In this case, the helicopters land and pick up passengers at a suitable spot outside the airbase, where airport regulations do not apply.
2. Due to budget cuts one airbase is active only three days a week. The area is often used for training purposes. If a pilot needs a precautionary landing on one of this airbase's inactive days, airport regulations state that the helicopter will not be able to take off after repairs are finished. A precautionary landing outside this airbase does not meet these restrictions, hence taking off is allowed.
3. Out on a training exercise a helicopter needed repairs. Since it was not stationed at home base, the equipment needed to comply with the regulations for 'working at heights' was not, and could not be made, available. The crew chose not to comply and did the repairs without being secured. This resulted in an operational helicopter and a successful training exercise.

2.2 Document study

We studied five incident/accident investigation reports and the formal responses of the Department of Defence to the House of Representatives. These investigation reports are:

1. Self-shooting: on 21 January 2019, an F-16, the follow-on aircraft in a formation of two, conducted a shooting session over the Vliehors. During the flight, the F-16 fired its on-board gun at a practice target (Inspectie Veiligheid Defensie, 2020b).
2. Dangerous goods: on 24 February 2016, officers from the Military Hazardous Substances Control Corps found an oxygen cylinder violation regarding its testing, labelling and release for road transport. The investigation committee made six recommendations. During a debate in the House of Representatives, a request was made to the government to commission an independent investigation into the implementation of the recommendations (Inspectie Veiligheid Defensie, 2020a).
3. Risks identified: from 17 to 19 June 2019, 14 students were on professional practice training at Woensdrecht Air Base and the nearby Ossendrecht military training ground. At 08.29 a.m. Wednesday morning, while preparing the assignment for the day, the group was unexpectedly caught in a thunderstorm. Some students were struck by lightning and one of them was seriously injured (Inspectie Veiligheid Defensie, 2021b).
4. Weak signals: the Defence Safety Inspectorate announced a study at the end of 2019 to gain insight into the handling of minor incidents at Defence and thereby strengthen the learning capability of the organisation. The research focuses on the function of signalling of minor incidents and the influencing factors (Inspectie Veiligheid Defensie, 2022).
5. NH-90 helicopter crash: on Sunday 19 July 2020, an NH-90 maritime helicopter with a four-man crew practised both approaches and deck landings on the patrol ship Zr. Ms Groningen. During the execution of a deck landing circuit, the helicopter hit the water. The two crew members in the cockpit did not survive the accident (Inspectie Veiligheid Defensie, 2021a).

These investigation reports are publicly available. The document study was enriched by the long experience of the first author (LB), who has spent years as an aviation psychologist working with and observing the population of interest.

2.3 Data analysis

LB conducted and transcribed the interviews, RB and VS read the transcriptions. LB used NVivo software to code and analyse the transcripts. The first level of analysis consisted of coding relevant sections line by line (Charmaz, 2011), resulting in categories of data called nodes. Using inductive open coding, nodes were created bottom up, emerging from the data rather than by applying a theoretical framework from the literature, although some nodes could refer to concepts found in the literature. Theoretical saturation was reached as the final three interviews generated no new information regarding our research questions (Saunders et al., 2018). The transcripts were coded using the previously created nodes. No new nodes were needed.

A total of 93 nodes were created. LB revised these twice to minimise overlap, to confirm a label accurately represented the findings and to make sure they contained a meaningful amount of data. This

work led to the removal of 27 nodes. A third revision combined the various nodes pertaining to the three example cases. Then, overly sensitive findings were excluded due to security reasons, resulting in the removal of another 13 nodes. This left 53 nodes for use in the analysis, which LB discussed with RB and VS.

On the second level of analysis, the co-authors (LB, RB and VS) identified relationships formed between the nodes, creating underlying themes that each contained several nodes. This analysis produced a total of 14 themes that seemed most pressing and/or promising regarding the difference between the way work is done and prescribed. The 14 themes were subsequently categorised in three overall themes used to describe the research.

As an example of the coding process, Figure 1 presents an excerpt of the transcription of interview 11. Nodes are set in {brackets} with the highlighted sentences that led to the node following the label. Three nodes {paper doesn't match context}, {rules don't match operational task} and {rules can't follow complexity/dynamics} resulted in the sub-theme of 'rules cannot keep up with the complex, dynamic nature' that subsequently became part of the overall theme 'need for adaptation in work'.

Q: Is it bad that there is a difference between the way work is done and prescribed?

A: [...]{**rules equal safety**} So from an MLA perspective, it is expected that you do follow the regulations because then you have implemented certain safety. {**gap not a problem**} In practice, I say, that's not a problem at all because it's about the intention of what do you want to achieve. [...] it has to do with the context {**paper doesn't match context**} and if the context, which is always changing, you will also sometimes have to take measures or implement things that also change over time. And which are then not on paper, or are on paper but are outdated or not applicable for that context. [...]Whereas someone who comes from university, I shouldn't say it like that, but someone who is purely in the rules, doesn't see the context, hasn't been operational, they look purely from the regulations. {**gap not a problem**} So the answer is, I personally don't mind. Even very good I think.

Q: You said that not one operator adheres to all regulations. Why is that?

A: {**rules don't match operational task**} Because those rules, they come from the civilian world and they are made up by people sitting behind a desk. And don't come from the real world. And of course there are a lot of rules...{**rules equal safety**} people want to write as many rules as possible because then it becomes "safer" [...].

Q: You just talked about context. Do you think regulations can follow the current dynamics and complexity of the operations?

A: {**rules can't follow complexity/dynamics**} No, no, we publish it today and tomorrow it's already outdated. And we certainly see that in the field of unmanned aviation. That explodes, well we really don't follow that. And we are now working on regulations and then it will be published next month, so to speak, it will already be outdated. [...] {**rules can't follow complexity/dynamics**} So no, so then you can indeed ask, regulation, safety, there will always be a delta in between.

Figure 1: A sample of a coded interview from interviewee 11 depicting the coded sentences (highlights) and accompanying nodes (in {brackets}).

Closely examining the data-based findings in two rounds of discussion, the co-authors debated the findings until consensus and compared the analysis with the literature.

3. Findings and interpretations

The findings are divided into three sections. The first two describe findings related to the importance of compliance and the need for adaptation in work. The third section, enduring tension, describes the findings related to the tension resulting from the contradiction between the importance of compliance and the need for adaptation.

3.1 Importance of compliance

The current study finds that adherence to rules and procedures is considered important in the population of interest. Rules are seen as the most important means to achieve safety and compliance. Each interviewee stated that compliance is motivated by the conviction that rules are important to safety:

> *I think rules are very important. They exist for a reason. We say rules are written in blood. (Interviewee 4)*

Senior management and rule makers/enforcers regard noncompliance (the difference between how work is done and prescribed) as undesirable because rules and compliance function as a safeguard:

> *A difference between an official regulation that says you're not allowed to fly when drunk on alcohol, for instance, and the way work is done is unacceptable. (Interviewee 1)*

These findings match the theory espoused by Rasmussen (Rasmussen, 1997), who states that the work space is defined by rules forthcoming from different domains, creating both safe and unsafe zones. Safety rules serve to define boundaries to prevent accidents (Amalberti et al., 2006). Rules and procedures enhance safety in some instances (Dekker, 2003a).

According to the interviewees, compliance in the RNLAF is also motivated by a fear of the consequences of noncompliance, meaning fear of personal liability and accountability.

> *We have these rule books with important rules, and if I break them, I will personally be held accountable, legally accountable. (Interviewee 2)*

The interviewees report that this fear of the consequences can lead to situations where safety concerns are diminished in favour of compliance.

> *So, I'm allowed to put the helicopter down on the other side of the airport fence, which is potentially far more dangerous because it's an uncontrolled area. […] But inside the airport fence, which is controlled terrain, we have rules that don't let [us put the helicopter down without sufficient firefighters]." (Interviewee 3)*

> *"If you return to base early because of a technical problem, you're not allowed to land. I'd have to deviate to an alternate [landing spot] until the fire department has finished monitoring the refuelling of other helicopters. […] But the airfield is right there! (Interviewee 2)*

In other words, compliance with rules and procedures is a way to manage liability and accountability of an organisation (Dekker, 2022).

Senior management reports compliance is an important goal in itself, particularly with regard to laws (rather than rules made by the organisation):

> *When it comes to compliance, it's about hard law, the bottom line. It doesn't mean you can't deviate, but [the deviation] has to be written down. (Interviewee 7)*

Additional data confirming that adherence to rules and procedures is considered important in the population of interest comes from the five incident investigation reports. These reports focus on how work *should* have been performed and how work-as-done deviated from how work is prescribed in rules and procedures. Differences between the way work is done and prescribed are in all five of these reports causally related to the occurrence. For instance:

> *When he was talking to air traffic control he declared an emergency but didn't make a PAN or mayday call and he didn't select the SSR 7700 code as prescribed by regulation. (document case study 1)*

> *Amidst perceived continuing pressure to validate her existence, the RNLAF is adapts her processes to the circumstances so that she can do with less. (document case study 5)*

People external to the organisation also see rules as the most important means to achieve safety and compliance. After an incident, the RNLAF promises the regulator, politics and public to implement new rules, procedures and instructions aimed at preventing a similar occurrence. With each new incident, the media questions this commitment and the resulting public debate puts pressure on the DoD to reconfirm and adhere to its own statements.

> *Why are things difficult? Because we live in the Netherlands, with crowded air space and politicians who will not accept that every now and then we'll drop an aircraft on a school playground. (Interviewee 12)*

This pressure can become so great that noncompliance is expected to have personal consequences such as the resignation for the minister of defence and chief of defence (Keultjes, 2017) ("Waarom Minister Hennis Moet Aftreden," 2017). Compliance, in other words, is a tool needed to control untrustworthy humans who inevitably make mistakes (Dekker, 2022).

From a leader's perspective, high levels of compliance (and therefore a small number of exceptions) are more manageable. Compliance with formal rules results in standardisation, transparency, predictability and control (Dekker, 2014c) (Weichbrodt, 2015):

> *The certificate means that the organisation is compliant with rules and so has assured safety to a certain extent. (Interviewee 11)*

The only side note rule makers/enforcers make on the ease of the compliance paradigm is the fact that knowledge of the rules is a prerequisite. However, rule makers/enforcers believe that knowledge of the existing rule base needs improvement.

> *No, we just put a person in [a] position because his career needs a post as commanding officer. Yes, of course, he's incredibly good at counting stuff, but he doesn't know [much] about military aviation regulations. (Interviewee 11)*

On the other hand, chances are that general knowledge of the rules has declined because so many rules are (too) detailed, (Laurence, 2005) or the notion that experience spreads the detailed knowledge of rules abundantly (Dahl, 2013).

3.2 Need for adaptation in work

According to the interviewees the number of rules is ever increasing, especially after an incident or accident. This matches the literature as advocated by Antonsen et al (Antonsen et al., 2008): new rules created after an occurrence provide an opportunity for an organisation to show they take safety seriously and try to prevent similar occurrences. The result is an increasing bureaucratisation of safety (Dekker, 2014c) (Dekker, 2022).

> *My brain is like an iceberg that holds lots of penguins. And with each new rule a new penguin appears. At some point the iceberg is full, no room for more penguin-rules. And as I get older, the iceberg starts to melt. In other words, with new rules, penguins fall off the iceberg. I'm just hoping it will be a penguin-rule about an area I need to avoid, because my map will show that too [to remind me], and not a the penguin-rule about the emergency procedure in case of an engine fire. (Interviewee 2)*

> *I don't think we should have a rule for everything. We just need a general set of rules that we can use to make the right decisions in situations we haven't [dealt with] before. But, take a look at our operations manual. […] As a pilot you can't know [all] these rules by heart. It's just too many, with 50 exceptions for every rule in the book. (Interviewee 5)*

Reducing the number of rules is seen as a necessity. However, some interviewees are very sceptical about the feasibility as they expect politicians and society will not accept the deletion of rules.

Besides signalling the plethora of rules, the current study has identified that the population of interest often finds it necessary to adapt rules in the execution of a task.

> *The world is changing much faster than rules of any kind can follow. We cannot even think that fast. (Interviewee 6)*

> *Every day we feel the tension between trying to fully comply with the rules and doing our job. (Interviewee 5)*

These are examples that no one thought of when writing the rules. (Interviewee 8)

These results match the research by Damoiseaux-Volman et al. (Damoiseaux-Volman et al., 2021a) which shows the difference between rules and the complex nature of work can be substantial. When comparing the way work is done to the way it is prescribed in guidelines for two tasks in a medical setting, they found ten and 16 new functions or changes, respectively,, indicating that work-as-done is far more complex than the rules convey.

Most interviewees were adamant: rules cannot keep up with the complex, dynamic nature of military operations in training situations or mission deployment. When it comes to less strict rules, most interviewees state that rules can never fully describe every situation encountered and that it is up to the operators to get the job done. And in some cases, it is permissible to deviate:

We have standard operating procedures, but if before a flight you decide to do things differently, that's fine, as long as you let your colleagues know. (Interviewee 7)

These results fit the work of Carim et al (Carim et al., 2016), who found that dealing with ill-suited procedures led pilots to use fragments of different checklists to cope with an emergency situation, therefore adapting the rules that could not follow the far more complicated situations they encountered.

The interviewees state that one reason rules cannot follow the complexity of daily work is that they are sometimes outdated and no longer match the context. Operators introduce a mismatch because they were not involved in writing the rules. Senior management as well as rule makers/enforcers agree that rules are written by people who do not perform the tasks themselves.

Policy makers at Central Staff who produce new rules and agreements with politicians stand too far away from the work actually done. They get so wrapped up in their own policy world and political climate that they have no idea what is needed at the operator level. (Interviewee 10)

These results go with the work by Carvalho et al (Carvalho et al., 2018), who found that rules written by officers who do not do the work themselves led to rules unsuitable for dealing with emergency response situations in firefighter organisations.

Every interviewee said that rules actively hamper task execution. According to operators, in some cases it would be simply easier and more convenient to not follow the rules.

I have a perfectly functioning aircraft out in the field after a precautionary landing which turned out to be a false alarm. Now I can't get the aircraft off the ground because there's a rule on it somewhere, that [even if] it's safe to go […] And that hampers task execution, because now the aircraft stays in the field unsheltered from weather, and maintenance and security [crew] need to [get to it] there. It just hampers us. (Interviewee 3)

Some interviewees found not being able to get the job done frustrating.

If we have to attend a forest fire [somewhere], I hope the rules will let us. That we won't have to ask [if] we are allowed to do this outside airbase opening hours […] or will we just put out the fire? (Interviewee 3)

As another example, in January 2022 the Minister of Defence stated in an interview that rules and procedures hamper defence against cyber-attacks and that laws and rules need to be changed ("Minister van Defensie: Procedures En Regels Bemoeilijken Aanpak Cyberaanvallen," 2022).

The literature confirms this perception. For instance, Tucker (Tucker, 2004) describes how disruption and errors in information hamper employees in doing their job. These 'operational failures' (Tucker, 2004) result in necessary workarounds (Halbesleben et al., 2008) (Tucker et al., 2020).

The need for adaptation in work is evident in that operational pressure takes precedence over compliance. Most interviewees report on air power receiving first priority, which makes sense to them, since this is this is the actual output of the RNLAF.

> *Mission first. We must have air power, that's the first priority […]. We're in a military world, which we sometimes forget. (Interviewee 11)*

Most interviewees refer to the drive to get the job done as can-do mentality. They see it as a positive aspect of the RNLAF identity, even though it can on occasion have a negative aspect.

> *The drive that people have to get it done, it's part of our DNA and undeniably has a lot of good in it […] but sometime a little dark edge to it. (Interviewee 9)*

As this quote demonstrates, the focus on operational output can result in a dark side to can-do behaviour. Especially rule makers/enforcers warn against this negative effect as they think can-do behaviour comes from employees looking for the loopholes.

> *We [they] are more 'can-do' than we dare realise. Of course, we [they] do 'plan-do-check-act', but we [they] also do 'can-do-react'. It takes us[them] a long way, but it can also make us [them] fall flat on our face. […] We [they] don't solve the problem, we [they] just try to figure out how we [they] can proceed as planned. (Interviewee 12)*

> *It's even about simple things, like driving regulations. Every truck driver is obliged to take a 30-minute break every three hours of driving. But when we [they] are on exercise in another country and need to move equipment, we [they] pay our drivers an exercise fee [..] so the rules don't apply to military drivers. (Interviewee 10)*

Rule makers/enforcers report that it is not always right to give first priority to air power, even though this is the desired output.

> *Not everything legal is smart to do. Not everything allowed is the sensible thing to do. I can get my neighbour pregnant, no law that prohibits that, but it's not smart. It'll give [me] so much trouble. (Interviewee 12)*

On the other hand, Morrison (2006) and Dahling et al (2012) have documented the positive side of noncompliance in favour of organisational goals, using the term 'pro-social rule breaking' to describe noncompliant behaviour benefitting the organisation, for instance, when employees need to choose between compliance or adaptability in the interest of task execution (Chung & Schneider, 2002). Letting operational tasks take precedence benefits the organisation and so this is considered constructive (Spreitzer & Sonenshein, 2004).

Finally, in the view of every single interviewee, changing rules that the RNLAF has *no* authority over is nearly impossible.

> *That's part of the rule base we, like every other ministry, have to comply with. […] There is no way to sugar-coat that. (Interviewee 9)*

> *The other day I was in a working group […] to see what is possible. Immediately someone said that it will be impossible to change the airbase into a stage field, like the Americans have. (Interviewee 1)*

Being able to change the airbase into a stage field means that after a precautionary landing, a pilot is allowed to take off when the necessary repairs have been done, even if the airbase is not active.

When it comes to changing rules that the RNLAF *does have* authority over, the process is not widely known, as the interviewees indicated.

> *I don't know if we have a procedure for adapting or changing rules. I hope so, but I'm not sure. I should know, but I don't. (Interviewee 8)*

Changing rules is a long, painstaking process. One reason is the lack of capacity for ongoing review and change, and as operators point out, the RNLAF imposes restrictions upon itself. For instance, applicability to all operators takes precedence over rules matching tasks, which requires exceptions. Senior management agree that it takes lots of effort to get a rule changed.

> *[It takes great force] to try to change the most important things, and that's nearly impossible. (Interviewee 5)*

According to half of the interviewees, it will need reassessment of the existing rule base and being able to adapt in work situations to resolve this tension.

> *We continuously need to assess our rule base with a plan-do-check-act loop. Are the rules still valid? (Interviewee 4)*

3.3 Enduring tension

The contradiction between commitment to rules and compliance on the one hand and the need for adaptation and freedom to use expertise on the other hand is apparent in an enduring tension within the RNLAF.

This tension is evident in examples where compliance prevails even if it leads to a riskier situation. For instance, because of airport regulations a certain number of fire fighters must be on hand when large aircraft with more than two passengers need to land or take off. Large helicopters, however, are allowed to land outside airports, without fire fighters present, due to a distinct set of rules. When not enough fire fighters are available, helicopters will land outside the perimeter of the airport to pick up passengers because the regulations prevent them from executing this task.

Being compliant does not necessarily mean being safe. […] The airport rules do not apply to areas outside airports, but then situation becomes less safe because you move from a bit of fire fighter capacity to none at all. (Interviewee 4)

The tension is also visible in examples in which compliance actually stifles operations. As interviewees explain, in some instances being compliant means the work just does not get done.

We have an exercise planned which requires me to have an emergency radio. So, I call the responsible department to ask for one. But I won't get this radio because it was not requested on the right form, even though they have them in stock. After filling out the proper form I do get the radio, but it hasn't been prepared. I have to load the settings myself, even though that's the job of the support department. The paperwork is more important than the operational tasks I have to do. (Interviewee 2)

These findings are substantiated by the literature on 'malicious compliance', which is defined as following the rules to the letter resulting in work not getting done or done unexpectedly (Dekker, 2018).

In other instances compliance can lead to unintended consequences, but because rules are adhered to and output is realised, according to the interviewees, unintended consequences do not get discussed. For instance, in the Netherlands, rules dictate how often a helicopter is allowed to land at designated helicopter landing spots:

When they get to that number, landing is no longer allowed. This rule is intended to minimise the sound burden for these areas as much as possible. However, the surroundings of these helicopter landing spots fall outside the extent of this rule. That means that helicopters are still allowed to land within as little as a couple of metres of a landing spot, because this falls under the rules of field landings. Obviously, this workaround makes it hard to meet the intent of minimising the sound burden. (Interviewee 10)

Becoming compliant with the intent of this rule would mean hampering task execution, so the unintended consequences get ignored and will not be discussed since literal compliance is realised.

Added to this is the limited interest senior management shows in the challenges posed by rules for daily operation. Senior management is aware of complex situations but shows no desire to understand or resolve the actual dynamics.

[When discussing complex situations with commanding officers] I have no idea why they don't. It's probably human behaviour, [them thinking] I've resolved the situation in the easiest way for myself, so why choose the hard way? (Interviewee 8)

Research by Tucker et al (Tucker et al., 2002) and Morrison (B. Morrison, 2015) seems to substantiate this finding. They found that the short-term success of problem-solving behaviours of operators diminishes the necessity for more structural change in the organisation.

Senior management sees change in the management of rules as a solution to the enduring tension. Accepting an ill-functioning rule is not an option:

If a rule hampers job execution, then together with rule makers/enforcers you try to find an alternative means of compliance that lets us continue our operation, or [find] a way to deviate

from the rules. But the answer can never be 'No that's impossible and that's the end of it.' Not changing a rule is a joint decision, with the operator involved. (Interviewee 6)

Considering the difficulties operators encounter when trying to change rules, it seems as if senior management underestimates how much effort it takes and how unsuccessful these attempts are.

The tension between committing to compliance and the need for adaptation is also seen in the way the RNLAF deals with instances in which operators use their own discretion to decide whether or not to be compliant. Of 12 interviewees, ten explicitly state that choosing to use your own discretion means justifying the course of action afterwards but having to explain their choice seems contradictory.

Operators don't cross safety lines […]. You won't risk your life on a simple training flight. (Interviewee 3)

If they make a decision and can explain it afterwards […] they will be supported. (Interviewee 9)

I am allowed to deviate if it means enhancing the safety of the flight. But I have to prove that safety was enhanced. But that is rather impossible, since I can't prove what would have happened had I not deviated. However, I have to explain in writing why I deviated from my flight authorisation. If [the explanation] is not accepted I have to go up one level and explain it to the next commanding officer. (Interviewee 2)

The interviewees signal a need for discretionary space, defined by basic rules, in which they can exercise their expertise and judgement as to how to execute the task. (Iszatt-White, 2007)

According to most interviewees common sense is vital to safety.

Undeniably, every now and then, including deployments, one has to land aircraft in unregulated places. […] We need to train [pilots] and give them the discretionary space to become and stay smart operators. (Interviewee 10)

However, it seems that in some cases, the more rules there are, the less common sense is used, until operators 'stop thinking' altogether.

I had this situation where pilots would fly back to the airbase when something was wrong with the aircraft, instead of making a precautionary landing as prescribed in the checklists. So, I addressed this situation and what I saw next was pilots always making a precautionary landing, even when flying back was the more sensible and practical course of action. In other words, they stopped thinking for themselves and simply followed what they saw was an order. (Interviewee 4)

This assertion is supported by Klein (Klein, 2011) who found that becoming comfortable with procedures leads to eroding expertise and the necessity to develop more skills seems to cease.

The tension between compliance and adaptation is exacerbated by the operators' lack of trust, which inhibits those involved from addressing the fact that the way work is done differs from the way work is prescribed. An explanation for this distrust is the perceived lack of a just culture:

> *It shows up the just culture, so to speak […] and it does harm, especially when [distrust] is directed at the employees involved. Personally, I think Central Staff and the Inspectorate seem to have an urge to jump at the chance to […] press charges, whatever. That doesn't help. (Interviewee 12)*

Research by Milliken et al. (2003), Tucker & Edmondson (2003) and Manapragada & Bruk-Lee (2016) finds that behaviour of management is related to keeping quiet about safety issues. Previous research shows that even closed accident investigations perceived as judgemental leads to the perception of diminishing social safety (Alta, 2022). Therefore, the way findings and conclusions are portrayed has an impact beyond the scope of an accident investigation (Boskeljon-Horst & Alta, 2023), possibly adding to the fear of consequences to being noncompliant.

A recent investigation (Boskeljon-Horst, Snoek, et al., 2023) showed that the RNLAF has a rather judgemental and retributive climate.

> *Every CO has a CO who also has a CO. Sometimes your CO simply demands you punish someone [after noncompliance], regardless of what you think the best response might be. Especially when something gets picked up by the media or politicians. Then the highest level of our organisation, the secretary and minister of defence exert pressure to respond retributively and this flows through the organisation down to our level. (Supervisor 1 in* (Boskeljon-Horst, Snoek, et al., 2023)*)*

> *Upper management uses what we call a long screwdriver to turn every CO in the hierarchy to the same position, so to speak. This is how they micromanage a situation deep in the capillaries of our organisation. It's a lack of trust and a refusal to let things go and be dealt with at the appropriate level (Supervisor 7 in* (Boskeljon-Horst, Snoek, et al., 2023)*).*

This same study (Boskeljon-Horst, Snoek, et al., 2023) showed that fear of personal consequences resulted in punishment for noncompliance, even if it was just for show.

> *His CO said to him: 'Okay, listen. Remember that incident last year? When the Commission of Inquiry comes, just bend over and take one for the team.' (Focus group 1 in* (Boskeljon-Horst, Snoek, et al., 2023)*)*

Previous research (Boskeljon-Horst et al., 2022) shows that if noncompliance is necessary to achieve goals, it is only visible if there is no personal risk involved or if it can be kept hidden from RNLAF senior management.

However, when it comes to safety, operators consider trust a priority. They have no intention to put themselves in danger by crossing safety lines. They do not need rules to tell them to stay safe.

> *That's one of the biggest issues in this organisation, the underlying problem, the lack of trust. Trusting the operator to make a decision. This professional knows best. Trust another, [trust] everything we do. (Interviewee 3)*

Rule makers/enforcers acknowledge that the organisation sometimes uses rules as a way to "cover its ass", enabling the appointment of blame if something goes wrong, a finding substantiated by Bye & Aalberg (Bye & Aalberg, 2020).

4. Discussion

The current study has identified that the population of interest considers adherence to rules and procedures important as our examples have shown. Rules are seen as vital to achieving safety and compliance. This is hardly surprising given the importance many organisations project onto rules (Hale & Borys, 2013a).

However, the existing rule base does not match the complex and dynamic nature of the work to be done, resulting in a need for adaptation of these rules. Since it is difficult to change rules, employees call on local ingenuity to achieve their goals despite the hampering rules (Boskeljon-Horst et al., 2022).

The public and political commitment to compliance that the DoD leadership makes reinforces a cycle of issuing promises followed by pressure to keep the promise. Formally acknowledging the hindrance of rules and reassessing, possibly deleting, some rules go against the promise to align the way work is done to the way work is prescribed. This limits the freedom of employees to act according to their judgement. The interviewees recognise the importance of rules and procedures. However, they are frequently confronted by rules that hamper task execution or reduce safety and that are hard to change. As research shows, this freedom, and in some cases accompanying non-adherence, may be necessary to get out of an unworkable situation (Klein, 2011). An accurate perception of this unworkable situation is required (Besnard & Greathead, 2003). The examples in the findings section show that operators are very well aware of the hampering aspect of rules and procedures and equally aware of the fact that, in some cases, solutions that adhere to the rules actually reduce safety. They recognise safer ways to perform, yet the rules do not allow this and changing them is too hard. Given the findings by Antonsen et al. (Antonsen et al., 2008), what is seen as a safety measure – rules and compliance – might become the opposite.

The crucial pressure on enforcing compliance on the one hand and the need for adaptation and freedom to use expertise on the other creates tension throughout the DoD. This tension is seen in malicious compliance resulting in riskier or stifled operations.

This tension is also seen in expecting an operator to justify, after the act, a deviation carried out for safety reasons and based on expert knowledge, which is a fallacy in itself. Presented as the freedom to deviate from rules in case safety is judged by the expert to be at risk, the expectation of justification shows this up as an illusion. If other experts, in hindsight, consider the deviation unnecessary, this use of expert knowledge is relabelled as a violation and the pressure to be compliant intensifies.

Finally, the tension is seen in the unwillingness to challenge the belief in compliance because there is no need for it. As a result of previously discovered local ingenuity (Boskeljon-Horst et al., 2022), the work gets done in spite of hampering rules. Because local ingenuity is non-transparent, the fact that operators have to choose between adherence and performing remains invisible. Hence there is no incentive to take on the difficult task of trying to change the existing rule base. Overcoming this lack of curiosity in how work gets done would obtain insight into the complexities of daily work as well as the difficulties employees encounter when they try to change rules. Because the belief in compliance is not challenged, operators are left with the tension between having to get the job done while hampered by rules and procedures.

These interpretations might explain the status quo of the differences between procedures and practice and the difficulties encountered when trying to adapt the existing rule base to match work situations. In the DoD's official safety narrative, compliance with rules is the (one and only) way to achieve safety. This contrasts with the need and desire for freedom, espoused not only by the interviewees but also in the literature (Besnard & Hollnagel, 2014) (Dekker, 2003a) (Hale & Borys, 2013a) (Klein, 2011), to exploit experience and expertise where rules provide insufficient support to the situation. However, the pressure to sustain this belief after embracing it so widely is simply too large to expect a shift in the compliance paradigm (Rae & Provan, 2019).

These findings and interpretations do not, however, seem to tell the whole story. Understanding the remaining of the gaps between work-as-done and work-as-prescribed requires an understanding of the differing interests and focus of the diverse levels. If these are too far apart, every effort undertaken to align procedures and practice is likely to fail.

Perceived at the top of the population of interest, the DoD is a political instrument that makes political promises that require political support. The press and budgetary debates are among its biggest threats. Regarding the latter, the DoD has suffered from a decade of budget cuts, while the political ambition and expected output have remained the same. Making sure nothing jeopardises the received budget seems a good motivation to stick to promises as much as possible. The DoD often acts as a social role model, as demonstrated by its recruitment advertising. This leads to a kind of moral superiority, expected by both politicians as society: soldiers are outstanding, moral, compliant citizens who represent an organisation that goes by the book. Compliance in itself is not that important, it is the promise of compliance that is vital.

For RNLAF senior management the interests are a little different. Adhering to the compliance narrative is important because they feel political pressure at their level as well. Breaking the political promise with noncompliance might have negative consequences, especially when the RNLAF needs to compete with other military forces for its annual budget. However, it is still expected to produce output (airpower). To achieve this, the RNLAF needs to be able to train-as-you-fight, preferably without the hindrance of rules and procedures. This means that, while appearing to be compliant, senior management needs either to condone noncompliance or not be curious as to how work is done, leading to a two-faced Janus situation; senior management expresses both the need for compliance as well as the need to do what is necessary to get the job done, without reflecting on or even recognising the contradiction. It seems that holding two contradicting views has become a subconscious routine. This results in the difference found between what is done within the RNLAF and what it promises the outside world: constant compliance.

At the operator level, the DoD/RNLAF is a military organisation with violence, i.e., airpower, as one of the most important means. Here the focus is on well-trained employees who can be deployed in any situation. Training-as-they-fight is the interest found at this level. Getting the job done needs adaptation, as seen in the local ingenuity adopted (Boskeljon-Horst et al., 2022). Compliance is important in situations where adhering to rules avoids serious physical risks for aircrew. More important, however, is avoiding personal liability or accountability. The fear of personal liability is aggravated by the perceived lack of a just culture. Combined, these findings provide a reason to refrain from transparency regarding the differences between practice and paper. Political support and the media are not the concern of operators. The division between budget and societal expectations are of some concern.

Every level of the RNLAF considers compliance important. However, our data suggests that, when going 'down' the hierarchy from political top to senior management to operator, there is a gradient in the explicit nature of our informants' willingness and ability to acknowledge this. Conceding that compliance is secondary to getting things done increases closer to operational levels. For senior management, the interest in retaining an image of a smaller distance between the way work is done and prescribed seems governed by political pressure; for operators "getting stuff done" (Dekker, 2018) may even be a source of professional pride of their workmanship (Boskeljon-Horst et al., 2022) (Deming, 1982) (Morel et al., 2008).

Visually depicting the emergence of differing interests (Figure 1) shows how the levels diverge in their focus.

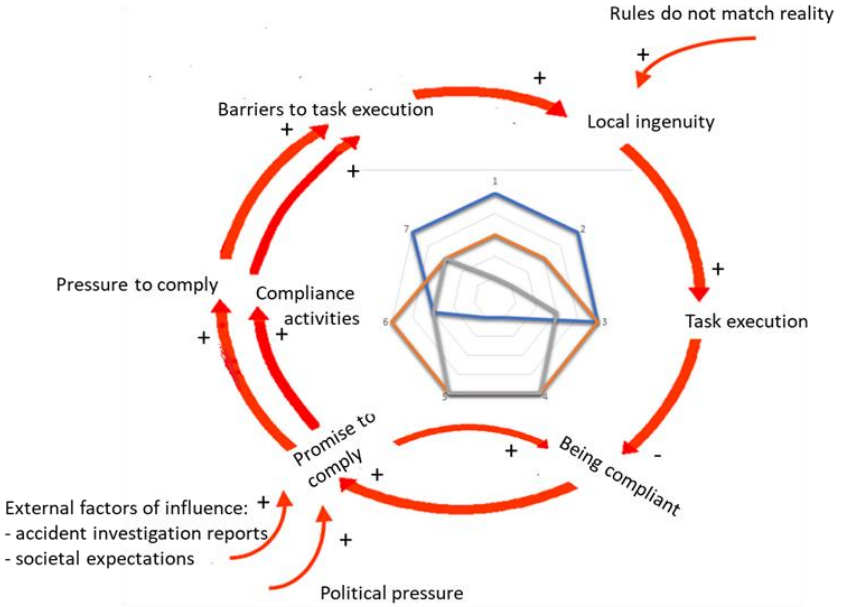

Figure 2: Different interests explaining the enduring differences between the way work is done and prescribed. The outer circle shows the emergence and reinforcing of the two contradictory views regarding compliance/adaptation. The spider diagram the middle shows the different interests of the political top in blue, the senior management of the RNLAF in orange and the operators in grey on seven themes: 1=political support; 2=press; 3=budget; 4=train-as-you-fight; 5=personal liability; 6=compliance; 7=societal expectations.

Figure 2 is created to model the differences in focus, showing a framework of interests. It is not underpinned by a quantitative analysis. The figure shows the need for adaptation demands local ingenuity in task execution, which means, however, less compliance. External evaluations combined with pressure from the political top results in the dual promise not only to be compliant but also to enhance compliance. This leads to more acts of compliance and increasing pressure to adhere to these promises, thus hindering task execution since rules and procedures often do not match the practice. Hence local ingenuity is needed again to get the job done. Looking at the seven themes arising from

this research – political support, press, budget, train-as-you-fight, personal liability, compliance, societal expectations – we can map the differing interests and focus in a spider diagram. As depicted, there is considerable difference between the focus of the levels, with RNLAF senior management functioning as a buffer between the political top and the operators with regard to political support and train-as-you-fight.

Reflecting on these findings and interpretations pinpoints a fear underscoring the interest in maintaining the difference between the way work is done and prescribed. The official safety narrative seems to convey a hidden message – bad things happen to bad people; in other words, if you do not obey (comply), you face bad things (personal consequences or accidents). This fear, arising from the threat of punishment, is reminiscent of a bogeyman, which provides an excellent metaphor. The idea of a bogeyman is eerily consistent with humanity's deference throughout the Common Era to some mythical functionary appointed to help govern the system, to ensure 'good behaviour', specifically to monitor and test human beings in it (Kelly, 2006).

As is commonly found in bogeyman stories, the personal consequences for being bad (not 100% compliant) remain unclear. In this case they can be deduced from the occasional sacrifice made to sustain the compliance belief, as seen in the resignation of both the minister of defence and the chief of defence (Keultjes, 2017) ("Waarom Minister Hennis Moet Aftreden," 2017) as well as lower down the organisation (Boskeljon-Horst, Snoek, et al., 2023). But these were personal choices. Choices by political and military command to resign and choices within the DoD to punish people. Unclear is what would have happened had those choices not been made, leading to the question of whether the bogeyman's fear of noncompliance is justified.

One opportunity to resolve the described tension is to embrace the pursuit of a "fifth-generation air force" (5GAF) to counter unrealistic expectations about compliance, and advance the use of expertise, common sense thinking and trust in operators. As Hale and Borys explain, the current aim to form a 5GAF contradicts the commitment to a compliance paradigm (Hale & Borys, 2013a). The focus on compliance leaves little room for using expert judgement and hence ownership and agility.

Realising a 5GAF might include a doctrine change, as investigated by Jahn (Jahn, 2019). He reports on a sudden doctrine change at the United States Forest Service where regulations changed overnight from prescriptive rules to guidelines. No extra rules were created, just a different interpretation of the existing rule base. The change gave operators more discretionary space and the professional judgement of the operators actually doing the work takes preference over rules made at a great distance from the place of work. Supportive modifications were also made to the investigation process and leadership training. Jahn's results (Jahn, 2019) are supported by research on the use of professional expertise by forest rangers [Kaufman, 1960 in (Dekker, 2022)] and sailors (Rochlin et al., 1987).

5. Conclusions

This research shows that—across multiple levels and roles—the RNLAF might desire to align work-as-done with work-as-prescribed and breach the status quo of gaps between the two. Its attempts to do so, however, will likely remain stymied because of a forceful dynamic balance of interests that both

produce and retain the status quo, while offering all stakeholders outwardly legitimate ways to elide its existence. We found that the higher-up leaders walk a tightrope, maintaining the opportunity for differences between the way work is done and prescribed, but agreeing superficially with the compliance focus and when undesired results become officially known, as with incidents and accidents. And although minor attempts are made to gain discretionary space and fewer rules, these are unsuccessful, because the pressure to keep the promise of compliance makes change unlikely. What we found was not only a self-sustaining force field of interests which simultaneously retains *and* denies noncompliance at almost all hierarchical levels we studied, but also the routine invocation of a nondescript organisational force-to-be-feared – reminiscent of a bogeyman - in case noncompliance was discovered. Efforts to solve the gaps between work as done and prescribed should be directed at this bogeyman: the belief system that compliance equals safety, adhered to by politics and public and given lip service by military leaders. By pressuring employees into compliance, we keep feeding the bogeyman, until the fear of freedom to exercise professional autonomy stifles the entire organisation. This study shows how the achievement of 5GAF behaviour and culture is hampered due to the enduring differences between the way work is done and prescribed. What could provide a solution is a change in doctrine so that rules as principles become rules as guidelines, embracing the ideals of the 5GAF. Further and deeper ethnographic research will be needed beyond the present paper to form a thorough understanding of the function and role of the 'bogeyman' in sustaining concurrent acquiescence and noncompliance.

6. Limitations and strengths

This study has both strength and limitations to consider. The use of a qualitative thematic analysis with a focus on exnovation enabled us to explain the consequences of two conflicting paradigms held in the same organisation. The qualitative methodology, using transcripts of semi-structured interviews and document analysis, provided a rich data set that shows the complexities and dynamics of everyday work hampered by the compliance paradigm. The official perspective of the DoD/RNLAF substantiated the perceptions of the interviewees.

Our research team, combining one researcher (LB) working for the RNLAF and external researchers, ensured a good mix of thorough knowledge of the DoD and RNLAF cultures and the tension between compliance and need for adaptation on the one hand and critical thinking on the other hand (Golden-Biddle & Locke, 1993). This ensured both the understanding and theoretical underpinning of the phenomena found.

However, the interviewee sample consisted of employees of a single organisation in the RNLAF and the chain of command following from the population of interest. Of concern are the sample size and representativeness of this sample. The sampling, however, was purposeful (Russell & Gregory, 2003). The sample size was deemed acceptable because no new information was provided, indicating a saturation point (Saunders et al., 2018).

The DoD and RNLAF face considerable external pressure to uphold the compliance paradigm. This means that the results of this study might not be generalisable to organisations in a different setting.

Nonetheless, although there is a significantly growing amount of research into the difference between the way work is done and prescribed, less is known about the reasons for maintaining these enduring differences, i.e., the various interests that reside in an organisation, and the resulting tensions. This study contributes to the theory by providing insight into these interests and stakes found in the Dutch DoD to retain these differences.

References

Alta, M. (2022). *Het effect van voorvalonderzoek op de beleving van sociale veiligheid van direct betrokken defensiemedewerkers*. Nederlandse Defensie Academie.

Amalberti, R., Vincent, C., Auroy, Y., & De Saint Maurice, G. (2006). Violations and migrations in health care: A framework for understanding and management. In *Quality and Safety in Health Care* (Vol. 15, Issue SUPPL. 1). https://doi.org/10.1136/qshc.2005.015982

Antonsen, S., Almklov, P., & Fenstad, J. (2008). Reducing the gap between procedures and practice - lessons from a succesful safety intervention. *Safety Science Monitor*, *12*(1). https://www.researchgate.net/publication/229005586

Antonsen, S., Skarholt, K., & Ringstad, A. J. (2012). The role of standardization in safety management - A case study of a major oil & gas company. *Safety Science*, *50*(10), 2001–2009. https://doi.org/10.1016/j.ssci.2011.11.001

Barrett, J. R. (2007). The researcher as instrument: learning to conduct qualitative research through analyzing and interpreting a choral rehearsal. *Music Education Research*, *9*(3), 417–433. https://doi.org/10.1080/14613800701587795

Besnard, D., & Greathead, D. (2003). A cognitive approach to safe violations. *Cognition, Technology & Work*, *5*(4), 272–282. https://doi.org/10.1007/s10111-003-0131-1

Besnard, D., & Hollnagel, E. (2014). I want to believe: Some myths about the management of industrial safety. *Cognition, Technology and Work*, *16*(1), 13–23. https://doi.org/10.1007/s10111-012-0237-4

Bieder, C., & Bourrier, M. (2013). *Trapping safety into rules: How desirable or avoidable is proceduralization?* Ashgate Publishing Co. https://doi.org/10.1080/00140139.2013.819675

Boskeljon-Horst, L., & Alta, M. (2023). Voorvalonderzoek contraproductief? De negatieve impact op sociale veiligheid en lerend vermogen. *Militaire Spectator*, *192*(3).

Boskeljon-Horst, L., Boer, R. J. De, Sillem, S., & Dekker, S. W. A. (2022). Goal Conflicts , Classical Management and Constructivism : How Operators Get Things Done. *Safety*, *8*(37).

Boskeljon-Horst, L., Snoek, A., & Baarle, E. van. (2023). Learning fom the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force. *Safety Science*, *161*.

Bye, R. J., & Aalberg, A. L. (2020). Why do they violate the procedures? – An exploratory study within the maritime transportation industry. *Safety Science*, *123*(November 2019), 104538. https://doi.org/10.1016/j.ssci.2019.104538

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. https://doi.org/10.1016/j.ssci.2016.06.008

Carvalho, P. V. R. d., Righi, A. W., Huber, G. J., Lemos, C. de F., Jatoba, A., & Gomes, J. O. (2018). Reflections on work as done (WAD) and work as imagined (WAI) in an emergency response organization: A study on firefighters training exercises. *Applied Ergonomics*, *68*, 28–41. https://doi.org/10.1016/j.apergo.2017.10.016

Charmaz, K. (2011). A Constructivist Grounded Theory Analysis of Losing and Regaining a Valued Self.

In *Five Ways of Doing Qualitative Analysis: Phenomenological Psychology, Grounded Theory, Discourse Analysis, Narrative Research, and Intuitive Inquirey.* The Guilford Press.

Chung, B. G., & Schneider, B. (2002). Serving multiple masters: Role conflict experienced by service employees. *Journal of Services Marketing*, *16*(1), 70–87. https://doi.org/10.1108/08876040210419424

Dahl, Ø. (2013). Safety compliance in a highly regulated environment: A case study of workers' knowledge of rules and procedures within the petroleum industry. *Safety Science*, *60*, 185–195. https://doi.org/10.1016/j.ssci.2013.07.020

Dahling, J. J., Chau, S. L., Mayer, D. M., & Gregory, J. B. (2012). Breaking rules for the right reasons? An investigation of pro-social rule breaking. *Journal of Organizational Behavior*, *33*(1), 21–42. https://doi.org/10.1002/job.730

Damoiseaux-Volman, B. A., Medlock, S., van der Eijk, M. D., Romijn, J. A., Abu-Hanna, A., & van der Velde, N. (2021). Falls and delirium in older inpatients: Work-as-imagined, work-as-done and preferences for clinical decision support systems. *Safety Science*, *142*(June 2020), 105355. https://doi.org/10.1016/j.ssci.2021.105355

Defensie. (n.d.). *5e generatie luchtmacht - wat houdt het in?* https://www.youtube.com/watch?v=5p9KXOrJbak

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

Dekker, S. W. A. (2014a). *Safety Differently. Human factors for a new era.* Apple Academic Press Inc.

Dekker, S. W. A. (2014b). The bureaucratization of safety. *Safety Science*, *70*, 348–357. https://doi.org/10.1016/j.ssci.2014.07.015

Dekker, S. W. A. (2018). *The Safety Anarchist*. Routledge Taylor & Francis Group.

Dekker, S. W. A. (2022). *Compliance Capitalism: How Free Markets Have Led to Unfree, Overregulated Workers*. Routledge Taylor & Francis Group.

Deming, W. E. (1982). *Out of the crisis*. MIT Press.

Golden-Biddle, K., & Locke, K. (1993). Appealing Work: An Investiation of How Ethnographic Texts Convince. *Organization Science*, *4*(4), 595–616.

Halbesleben, J. . B., Wakefield, D. ., & Wakefield, B. . (2008). Work-arounds in health care settings: Literature review and research agenda. *Health Care Management Review*, *January-Ma*.

Hale, A., & Borys, D. (2013). Working to rule, or working safely? Part 1: A state of the art review. *Safety Science*, *55*, 207–221. https://doi.org/10.1016/j.ssci.2012.05.011

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. *Paper Presented at the Gestaltung Nachhaltiger Arbeitssysteme, 58. Kongress Der Gesellschaft Für Arbeitswissenschaft, Pp.19-24.*

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Hutchins, E. L. (1995). *Cognition in the wild.* MIT Press.

Inspectie Veiligheid Defensie. (2020a). *Veilig over land en door de lucht. Onderzoek naar maatregelen voor het vervoer van gevaarlijke stoffen door Defensie.*

Inspectie Veiligheid Defensie. (2020b). *Veiligheid in de lucht. Onderzoek naar een zelfbeschieting van een F-16 boven de Vliehors.*

Inspectie Veiligheid Defensie. (2021a). *Luchtvaartongeval NH-90. Onderzoek naar luchtvaartongeval NH-90 Aruba 19 juli 2020.*

Inspectie Veiligheid Defensie. (2021b). *Risico's onderkend? Onderzoek naar een blikseminslag op oefenterrein Ossendrecht 19 juni 2019*.

Inspectie Veiligheid Defensie. (2022). *Sterker reageren op zwakke signalen. Onderzoek naar de signaalfunctie van lichte voorvallen.*

Iszatt-White, M. (2007). Catching them at it: An ethnography of rule violation. *Ethnography*, *8*(4), 445–465. https://doi.org/10.1177/1466138107083562

Jahn, J. L. S. (2019). Shifting the safety rules paradigm: Introducing doctrine to US wildland firefighting operations. *Safety Science*, *115*(January), 237–246. https://doi.org/10.1016/j.ssci.2019.02.002

Kelly, H. A. (2006). *Satan: a biography*. Cambridge University Press.

Keultjes, H. (2017). Minister Hennis treedt af om fataal mortierongeval Mali. *Het Parool*. https://www.parool.nl/nieuws/minister-hennis-treedt-af-om-fataal-mortierongeval-mali~b7525e0f/?referrer=https%3A%2F%2Fwww.bing.com%2F

Klein, G. (2011). *Streetlight and Shadows: Searching for the Keys to Adaptive Decision Making*. MIT Press.

Laurence, D. (2005). Safety rules and regulations on mine sites - The problem and a solution. *Journal of Safety Research*, *36*(1), 39–50. https://doi.org/10.1016/j.jsr.2004.11.004

Lawton, R. (1998). Not working to rule: Understanding procedural violations at work. *Safety Science*, *28*(2), 77–95. https://doi.org/10.1016/S0925-7535(97)00073-8

Manapragada, A., & Bruk-Lee, V. (2016). Staying silent about safety issues: Conceptualizing and measuring safety silence motives. *Accident Analysis and Prevention*, *91*, 144–156. https://doi.org/10.1016/j.aap.2016.02.014

McDonald, N., Corrigan, S., & Ward, M. (2002). Well-intentioned people in dysfunctional systems. *Paper Presented at the 5th Workshop on Human Error, Safety and Systems Development*.

Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of Management Studies*, *40*(6), 1453–1476. https://doi.org/10.1111/1467-6486.00387

Minister van Defensie: procedures en regels bemoeilijken aanpak cyberaanvallen. (2022). *Informatiebeveiliging Nederland*. https://informatiebeveiliging.nl/minister-van-defensie-procedures-en-regels-bemoeilijken-aanpak-cyberaanvallen/

Moppett, I. K., & Shorrock, S. T. (2018). Working out wrong-side blocks. *Anaesthesia*, *73*(4), 407–420. https://doi.org/10.1111/anae.14165

Morel, 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. https://doi.org/10.1518/001872008X250683

Morrison, B. (2015). The problem with workarounds is that they work: The persistence of resource shortages. *Journal of Operations Management*, *39–40*, 79–91. https://doi.org/10.1016/j.jom.2015.07.008

Morrison, E. W. (2006). Doing the job well: An investigation of pro-social rule breaking. *Journal of Management*, *32*(1), 5–28. https://doi.org/10.1177/0149206305277790

Rae, A., & Provan, D. (2019). Safety work versus the safety of work. *Safety Science*, *111*, 119–127. https://doi.org/10.1016/j.ssci.2018.07.001

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science*, *27*(2), 183–213. https://doi.org/10.16250/j.32.1374.2016270

Reason, J. (1990). *Human Error*. Cambridge University Press.

Reason, J. (2008). *The Human Contribution. Unsafe acts, accidents and heroic recoveries.* Ashgate Publishing Limited.

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. http://www.caso-db.uvek.admin.ch/Documents/MTh_SM_22.pdf%5Cnpapers3://publication/uuid/62B6CC6F-7E71-4540-94CB-32E1575D24CA

Russell, C. K., & Gregory, D. M. (2003). EBN users' guide: Evaluation of qualitative research studies Clinical scenario. *Evidence-Based Nursing*, *6*(2), 36–40. http://ebn.bmj.com/

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality and Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Soh, S. L., Lane, J., & Tan, C. (2020). Researcher as instrument: a critical reflection using nominal group technique for content development of a new patient-reported outcome measure. *International Practice Development Journal*, *10*(2), 1–9. https://doi.org/10.19043/ipdj.102.010

Spreitzer, G. M., & Sonenshein, S. (2004). Toward the Construct Definition of Positive Deviance. *American Behavioral Scientist*, *47*(6), 828–847. https://doi.org/10.1177/0002764203260212

The Hague: Onderzoeksraad voor de Veiligheid. (2018). *Draadaanvaring Apache-helikopter tijdens nachtvliegen (Wire strike during Apache-Helicopter night flight)*.

Tucker, A. L. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management*, *22*(2), 151–169. https://doi.org/10.1016/j.jom.2003.12.006

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures: Organizational and Psychological Dynamics that Inhibit System Change. *California Management Review*, *45*(2), 54–72.

Tucker, A. L., Edmondson, A. C., & Spear, S. (2002). When problem solving prevents organizational learning. *Journal of Organizational Change Management*, *15*(2), 122–137. https://doi.org/10.1108/09534810210423008

Tucker, A. L., Zheng, S., Gardner, J. W., & Bohn, R. E. (2020). When do workarounds help or hurt

patient outcomes? The moderating role of operational failures. *Journal of Operations Management*, *66*(1–2), 67–90. https://doi.org/10.1002/joom.1015

Waarom minister Hennis moet aftreden. (2017). *Trouw*. https://www.trouw.nl/opinie/waarom-minister-hennis-moet-aftreden~b17626f0/?utm_source=link&utm_medium=app&utm_campaign=shared content&utm_content=free

Weichbrodt, J. (2015). Safety rules as instruments for organizational control, coordination and knowledge: Implications for rules management. *Safety Science*, *80*, 221–232. https://doi.org/10.1016/j.ssci.2015.07.031

Woods, D. D., & Shattuck, L. G. (2000). Distant Supervision-Local Action Given the Potential for Surprise. *Cognition, Technology and Work*, *2*(4), 242–245.

Yoon, B., & Uliassi, C. (2022). "Researcher-As-Instrument" in Qualitative Research: The Complexities of the Educational Researcher's Identities. *Qualitative Report*, *27*(4), 1088–1102. https://doi.org/10.46743/2160-3715/2022.5074

Chapter 5. Learning from the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force

Leonie Boskeljon-Horst, Anke Snoek, Eva van Baarle (2023)

Safety Science 161 (May)

Doi: 10.1016/j.ssci.2023.106074

Learning from the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force

Abstract

In any organization, a restorative just culture and learning from incidents are key to improving organizational safety. The focus of a restorative just culture is not on determining right or wrong, but on repairing (restoring) and learning from an occurrence, by bringing both culprits and the people affected by the occurrence together to collectively answer the question of what to do next. However, a competitive and hierarchical organizational culture can make such openness difficult. Little can be found in the literature on the complexities of fostering a restorative just culture in a predominantly retributive setting. This article presents a qualitative case study aimed at understanding the complexities of fostering a restorative just culture in practice in a retributive setting. The article describes a safety standdown in the Royal Netherlands Air Force, based on observations, focus groups and semi-structured interviews. The findings reveal a spontaneous restorative response but an actual restorative setting was only partly attained. An inductive, thematic analysis of the data identified four key themes: 1) the need for vulnerability; 2) moral courage of leaders, 3) resolving the tension between learning and punishing, and 4) preventing a hybrid form of (un)just culture (asking people to openly share information, and then punishing them for doing so). The analysis shows that a hierarchy that appreciates and encourages a retributive response complicates the endeavour to create a restorative just culture. Upper management must deal productively with these challenges to foster a restorative just culture.

1. Introduction

Learning from accidents, incidents or near misses allows organizations to improve safety (Dekker, 2006) (Reason, 1997) (Woods et al., 2010). This is particularly key to organizations operating in high-risk environments (Catino & Patriotta, 2013). To learn and improve, information about these occurrences is vital, particularly the information that employees involved in an occurrence possess. How an organization responds to an occurrence determines the quantity and quality of information that runs freely throughout the organization. It determines how safe people feel to share their concerns.

One way to invest in freely flowing information is by fostering a just culture, a concept introduced by James Reason (1997) and embraced in such sectors as aviation, healthcare, nuclear organizations and the railway (Heraghty et al., 2021). Just culture entails a compromise between a fear of blame-free approaches and the need for accountability on the one hand and the willingness of employees to voluntarily report safety-related issues without being punished for it on the other hand (Dekker & Breakey, 2016; International Civil Aviation Organization, 2018).

Often just culture is seen as an atmosphere of trust, wherein events are shared and reported (Kováčová et al., 2019) and there is clarity on what is and is not accepted (EUROCONTROL, 2006; Balcerzak, 2017). In contrast, models developed to determine whether an event is classified as a mistake or as negligence and which draw a clear line between acceptable and non-acceptable behaviour (Reason, 1997) (Marx, 1997) lead to a retributive just culture (Dekker, 2016). Retribution is deeply engrained in such domains as aviation and healthcare and has a long tradition in the hierarchical and competitive culture of many organizations (Dekker, 2011). The retributive approach is a fairly quick, clear-cut process. The consequences of errors are clear to everyone (International Civil Aviation Organization, 2018), acceptable and non-acceptable behaviour are clearly defined (Reason, 1997) and balance is restored by holding someone accountable (Dekker, 2016). As long as the consequences are proportionate to the actions, retribution is seen as just, resulting in retributive justice (Dekker, 2016).

However, retributive measures that are perceived as unjust can have a detrimental effect on the open culture one tries to create (Reason, 1997; Lipira & Gallagher, 2014; Lawrenson & Braithwaite, 2018; Brborović et al., 2019). Acceptable and unacceptable behaviour are not fixed categories but are determined by time, context and the person who makes the judgement (Dekker, 2003; Alicke et al., 2008; Dekker, 2009). When a judgement seems unjust, people might try to hide their mistakes to avoid negative consequences (Dekker, 2007e) and the willingness to report errors or occurrences will drop. As Ivan Pupulidy put it elegantly, “The currency of safety is information” (Safety Monkey podcast, episode 11, 2020).

To counter the negative effects of retribution, a paradigm called restorative just culture is gaining popularity within organizations. In a restorative just culture the focus is not on determining whether behaviour is right or wrong, but on repairing (restoring) and learning (Dekker, 2016; Dekker & Breakey, 2016; Goodstein & Aquino, 2010). The conversation between the ‘culprits’ and the people impacted by the occurrence is forward-looking with the aim of understanding what happened, gaining knowledge on the context, system and part of the organization involved and the question of what to do next (Parkinson & Roche, 2004). This does not mean that restorative just culture lets people ‘get away with stuff’. On the contrary. People are still held accountable, but instead of focussing on punishment imposed by an impartial party which in the case of severe accidents and crimes is hardly ever satisfactory for the victims (Braithwaite, 2002), restorative justice focusses on what needs to be done to repair trust and relationships. By involving everyone affected by an event this becomes a broadly-based decision; the offender participates in deciding what needs to be done, including the consequences for the offender. Research shows that even with major events such as sabotage or vicious crimes, the restorative approach has more satisfying results than the retributive approach for the victims, the offenders and communities (Braithwaite, 2002).

Restorative just culture appears to have a positive impact on safety culture in organizations. Turner et al. (2020) describe how implementing a restorative just culture increased trust, shared accountability and learning. A focus on repairing and learning is found to result in both quantitatively and qualitatively richer information to detect near misses, system solutions and a safer workplace (Dekker & Breakey, 2016; Bell et al., 2011). Kaur et al. (2019) found that introducing a restorative culture seems to lead to fewer suspensions and disciplinary measures, a reduction in absence due to illness, and increased disclosure and cost effectiveness.

Although these studies illustrate the positive impact of restorative just culture, not much has been written on the complexities of developing this culture in an organization accustomed to retributive just culture. As retributive measures can have a damaging effect on the open culture one tries to create (Lawrenson & Braithwaite, 2018), understanding the complexities arising from fostering a restorative just culture in a retributive just culture context is vital. Therefore, the objective of this study is to identify and understand the complexities that arise when attempting to foster a restorative just culture in a mostly retributive organization. This paper presents a detailed account of a situation in the Royal Netherlands Air Force in which efforts to resolve the tension between the retributive response and restorative approach only partly succeeded. It identifies both obstacles and success factors.

The following describes the research setting and the methods and then presents the findings on the impediments and tensions that arise when establishing a restorative just culture. This study concludes with a discussion of the findings.

2. Method

2.1 Research setting

In 2019 the authors of this paper took part in an action research project aimed at fostering a restorative just culture in the Dutch Department of Defence. During the data collection phase, the authors came across a particular situation that could serve as an example of to how to foster a restorative just culture. The case study method (Yin, 2003) was applied to understand the complexities that arise when fostering a restorative just culture in a retributive context.

The situation in question was a 'safety standdown,' held in the Royal Netherlands Air Force (RNLAF). A safety standdown is a RNLAF intervention that cancels all flights and brings the aviation community of a squadron together to discuss the safety matters at stake. A safety standdown usually happens after a serious incident and can last from a couple of hours up to a full day depending on the number of matters that need to be discussed.

2.2 Data collection

Data for the case study was collected in two rounds. The first was an explorative phase, in which five individual and four focus group interviews were conducted to determine which themes regarding social safety and just culture are important for the RNLAF. These interviews were conducted by the first author (LB) and the second author (AS). Previously employed by the RNLAF as an aviation psychologist, LB is familiar with this squadron. As an outside researcher, AS is unfamiliar with the RNLAF and this squadron. Participants in the focus groups were all pilots or crew members and all volunteered to take part in this research. Individual interviews were conducted with senior officers who are also pilots. They were invited because they are the commanding officers of the participants in the focus groups. These interviews were intended to obtain a clear picture of just culture in the RNLAF domain. Most respondents

mentioned the safety standdown because of the unique restorative dynamics they had observed between the 'offender' and his colleagues.

In the second round, another four individual interviews were conducted to elaborate on the mentioned safety standdown. This article draws on observations by the first author (LB) who attended the particular safety standdown and detailed accounts of key actors: 1) the pilot who presented his case during the safety standdown, 2) a crew member who presented his own case in the same safety standdown, 3) the commanding officer (CO) who organized the event and 4) the aviation psychologist and first author who attended this safety standdown. These individuals are referred to as case study critical interview interviewee #.

A total of 13 interviews and focus groups were conducted, involving 29 employees. Focus group interviews lasted approximately two hours and one hour for an individual interview. Interviews were guided by a topic list, with questions on the dynamics between restorative and retributive responses to incidents in the RNLAF. Interview questions centred on how the involved parties experienced the safety standdown. The interview schedule covered five types of questions: 1) which events led to the safety standdown, how did these events influence them as individuals, and how did they affect the group; 2) how did they prepare for the safety standdown, which choices did they make; 3) how did they experience the safety standdown in terms of both presentations and the response of the group; 4) what were the long term effects of the safety standdown; and finally 5) how did they experience the tension between restorative and retributive tendencies? Topics were added to the list when new themes emerged. The interviews were conducted until data saturation was reached, meaning when no novel themes emerged after participant interviews (Saunders et al., 2018).

2.3 Data analysis

The data collected in the explorative phase were transcribed verbatim and analysed with NVivo software. The first level of analysis consisted of a rigorous round of open coding. Relevant sections were coded line by line (Charmaz, 2011). Nodes could refer to concepts in the literature but were mostly created bottom up, emerging from the data rather than by applying a theoretical framework from the literature. As the interviews proceeded, the coding became more fine-grained and comparative. At the second level of analysis, axial codes were developed to outline the relationships between the nodes. This analysis resulted in several positive and negative themes that seemed most pressing in the RNLAF and subsequently contributed to the case study.

After the second round of data, transcripts and observations were analysed independently by the researchers to ensure inter-rater reliability and the findings were extensively examined in two discussion rounds.

The research was conducted and analysed by two insider researchers who work for the RNLAF (LB) and the DoD (EvB), and one outsider researcher (AS) to ensure a good mix of thorough knowledge of the DoD culture and critical thinking.

3. Findings from the restorative safety standdown in the RNLAF

This section first presents the safety standdown, drawing on the interviews and observations. It then analyses the complexities of trying to foster a restorative just culture in a retributive setting.

3.1 Background and set-up of the safety standdown

In a training mission overseas, pilot X made judgement calls in three separate events that, combined, made his superiors doubt his capabilities. The CO of the squadron discussed the situation with his own chain of command and asked for trust to deal with the matter on his own. *"He's my guy, let me handle this"*. The trust and leeway were provided. The CO decided to repatriate the pilot. In line with the principles of restorative just culture, the CO and other superiors sat down with pilot X, raised their safety concerns, and invited him to share his side of the story. His superiors were satisfied with pilot X's explanations, remorse and sense of responsibility, and pilot X received no official warning.

In the months following, Pilot X noticed that his repatriation was the source of gossip.

> *After I came back, some of my friends called me to ask what happened. (…) But many other people had already formed an opinion and didn't want to ask about the incident. People started doubting me, and my flight safety. Some people said, "This [the gossip about his capabilities] is all bullshit," but others said, "Why does he still have his wing[3]?" (Pilot X)*

The CO noticed that the community were unsure whether the repatriation was a punishment or not. Motivated especially by peer concerns about pilot X's ability to fly safely, the CO decided together with pilot X to hold a safety standdown to provide clarity on how the occurrences had been handled, hopefully end all the gossip, and importantly, restore trust among the colleagues.

> *So, we decided to explain things in a safety standdown (…) to calm down the emotional responses. I think it did a lot of good for the organization and for me (…) even if it might have come a bit late [the safety standdown happened nine months after the incident]. (Pilot X)*

The intention of the safety standdown is to learn from the occurrences, in line with just culture, and at the same time draw a metaphorical line between acceptable and unacceptable behaviour. This safety standdown was attended by the aviation section of the squadron, about 40 of the total 50 people. The safety standdown began at 0900 and ended around 1630.

To make this safety standdown less confronting for pilot X, they added a presentation on another event involving a crew member. Both speakers would talk about their events and the rest of the aviation community would be able to ask questions. The safety standdown was conducted in a building used for lectures and briefings where the presenter stands on the podium and the audience sit in ascending rows of seats much like in a movie theatre.

[3] This is a badge in the shape of a wing that a pilot wears on the uniform. It is the visual representation of a pilot license and being a pilot.

3.2 First presentation: wrong material

The crew member's event was relatively minor compared to pilot X's. Just before a training mission, he discovered that he had been given the wrong material. The crew member concluded that there was no extra risk involved and weighed the impact of the wrong material against the necessity of the mission. Aborting this training mission would lead to the abortion of consecutive missions as well. Instead of discussing the matter with his supervisor, the crew member decided to go ahead with the mission and use the material provided.

Standing behind a lectern on the podium, the speaker used PowerPoint slides to present a technical, chronological story about his risk decision process.

During Q&A, the questions were accusing, and the audience simplified the speaker's complex situation to a seemingly rhetorical dichotomy: *"Did you not know you were breaking the rules, or did you just not care?"* Taken aback, the speaker moved away from the audience. In the RNLAF it is quite common to address a colleague bluntly under the precept that it 'isn't personal.' However, the questions did feel personal to the speaker because they oversimplified his complex situation.

The speaker recalls this experience as something he was not accustomed to. In his previous squadron, he was used to giving full disclosure during safety standdowns, and this was his intention during this safety standdown. Contrary to his expectations, his story was met not with understanding but with negative judgement.

> *I did a PowerPoint presentation with a timeline to show exactly what happened, thinking it would lead to a better understanding of my situation. But I had to face some upsetting comments, which made me feel cornered. A few colleagues approached me afterwards, saying they would now think twice if asked to share a story in a safety standdown. (Case study interviewee 2)*

3.3 Second presentation: risky judgement calls

Pilot X had, unexpectedly, quite a different experience. He said that he had been nervous preparing his presentation but had chosen the vulnerability approach: to not focus on the practicalities of the incident, but on what it had done to him as a person. His goal was to make the audience understand why his choices made sense to him at the time, how he reflected upon them in hindsight, and what he had learned in the process.

> *My aim was to be completely open because I still stand by my decisions. In hindsight I realise I should have been more open to alternatives. (…) Was I wrong? Absolutely. (…) I could have picked [another solution]. (…) I really wanted to show what caused the situation and how my character played a role in the choices I took in the circumstances. (…) I felt strongly that people deserved an honest story. (Pilot X)*

The presentation style that fitted his goal was an honest narrative. He decided not to hide behind the lectern, using PowerPoint slides, but to stand right in front of the audience. Rather than discussing the facts of the occurrences, he told a story about himself: who he is, what motivates him, what his

weaknesses and strengths are, his pitfalls and how difficult it was to stand in front of his colleagues disclosing all this personal information but, simultaneously, how important it was for him to help the audience understand his personal motivation and the challenges he encounters in his job. He outlined how his can-do mentality that had been instilled in him during his training as a marine contributed to his deviating from certain safety measures.

> *I did those things not because I wanted to screw things up but because (…) I can be very goal- and task-oriented. (Pilot X)*

He also made a public promise to do all he could to keep everyone on board his helicopter safe.

The aviation psychologist observed that during pilot X's hour-long talk, the audience was completely silent. The audience witnessed a sensitive colleague making an honest effort to share his personal feelings, deliberations, and vulnerability. Not excusing himself but explaining and apologising if his behaviour had caused doubts or concerns. And when it was time to ask questions, there were hardly any, mostly comments that showed understanding, compassion, and acknowledgement.

Pilot X said that after his presentation, many people who had judged him badly before now came up and said that they felt confident of his capabilities again and thanked him for his honesty. The positive responses exceeded his expectations.

> *I'm really happy about how it all went. (…) It's never fun to admit to a group of your peers that you made a mistake. (…) I would've liked none of this to happen, that everyone would have said, "Fantastic, well done, zero mistakes". (…) I don't think that this experience has made me personally better, but as a community, we have grown because of it. (…) So, yes, on the one hand, it was a lesson in humility but on the other hand, being completely honest – like "this is what I did, please don't do it like I did" – helped me find my place in the group again, and in myself. (…) It was really good to be able to tell my side of the story.*

Pilot X said it was a painful experience, but also liberating. He often uses this example to train others.

> *[The message is not] "don't be stupid enough to make this kind of mistake," but "this is how we deal with mistakes, and it's okay to make mistakes."*

Pilot X recalled the experience as something that did both the RNLAF and himself a lot of good. As hoped, the judgemental gossiping about pilot X ceased after the safety standdown. He received no formal punishment for the three events. His performance at the safety standdown was seen as the accountability he needed to provide to restore the relationships. Nowadays, pilot X is a commanding officer himself. He admitted that his repatriation and the subsequent gossip first made him doubt he had a further career at the RNLAF, but that the safety standdown had restored his trust. He is glad no punitive measures were taken:

> *I think that if you punish someone who makes mistakes with the best of intentions (…) it would be counterproductive. (Pilot X)*

4. Fostering a restorative just culture in a retributive setting

This case study reveals four themes key to fostering a restorative just culture in a retributive setting: 1) The need for vulnerability; 2) Moral courage of leaders; 3) The tension between learning and punishing; and 4) A restorative response in a retributive context.

4.1 The need for vulnerability:

Standing in front of one's peers, confessing one's mistakes, makes speakers vulnerable, which makes it understandable if they detach themselves somewhat from their presentation. As described above, the first presentation in the RNLAF safety standdown was technical, supported by PowerPoint slides. The speaker did not reveal what he personally felt about the incident. His presentation elicited a retributive response from the audience, as indicated by their dichotomous, closed, rhetorical questions that focussed on what had happened, not why it had happened, their lack of empathy considering the challenges the speaker had faced, and the speaker feeling that he was under attack.

The second presentation was a personal, emotional narrative covering far more than the incidents. The response it elicited was restorative, as indicated by the expressed appreciation for the speaker's honesty, the lack of questions on what had happened, the acknowledgement of the challenges faced and a recognition that these three incidents could happen to anyone. The shift in the audience's response was unprompted; it happened spontaneously.

The pilot of the second presentation narrates how he experienced the safety standdown:

> *On the one hand, it felt liberating. I felt that this was needed. I felt strongly that people deserved an honest story of what happened. (…) On the other hand, it felt very painful, and I felt vulnerable. (…) I have learned that by expressing my vulnerability, people start to understand you better as a person. (Pilot X)*

Other respondents shared this view:

> *I think that being honest, expressing one's vulnerability, always works. I think that is the best way to communicate about stuff like these [fuckups]. (Focus group 2)*

> *If you are open and honest towards others, then people dare to share confidential things with you as well. (Focus group 3)*

However, within the RNLAF, the culture is not always safe enough to express such vulnerability, even in the setting of the safety standdown. At the RNLAF it is customary practice to be singled out in front of the group when sharing an incident during a safety standdown. The interviews revealed that in other squadrons participants sometimes make disparaging monkey sounds during a safety standdown.

> *It is an instant killer for this whole culture when people react like that. You might as well start all over again [establishing a just culture], because you have completely missed the goal. (Focus group 1)*

The retributive response, responding with blame to openness, as found in this study of the safety standdown, also arises in other RNLAF traditions.

> *We have this tradition that when you screw up, you pay a fine in the form of beer. That's how we collect enough beer for our squadron parties. You can get fined for a stupid remark or an error, even if it's an honest one. Then, you have to stand on a chair in front of all your colleagues and say who you are and what you did. It's supposed to be a joke, but people feel punished. It's the price you pay for being open and honest. (Supervisor 3)*

4.2 Moral courage of leaders

The second theme identified is leadership, not just in the squadron but in the whole chain of command, and specifically the moral courage leaders need to act restoratively.

The interviews showed that the chain of command was worried that pilot X was going rogue when he made the risky judgement calls. However, his CO explicitly asked his superiors to be given the freedom to handle the situation in a restorative matter and he was granted this trust. Respondents described how important it is for them to have a superior that supports them and grant them their trust when occurrences happen. Other COs confirmed that often there is pressure from above to punish. Not giving in to this pressure can have a detrimental impact on their career and promotion chances because being restorative can seem to be a weakness, an inability to deal with the situation.

> *I've reached my final rank, so I don't have to worry about not getting promoted. That's quite liberating. (Case study critical interview interviewee 3)*

COs felt they had a mandate to deal restoratively with less visible occurrences. But when something happens that might cause a media storm or political turmoil, such as questions from the House of Representatives, COs felt pressured by upper management to act retributively. One CO explained:

> *Every CO has a CO who also has a CO. Sometimes your CO simply demands you punish someone, regardless of your own opinion of what the best response may be. Especially when something gets picked up by the media or politicians. Then the highest level of our organization, the secretary and minister of defence exert pressure to respond retributively, and this flows down through the organization to our level. (Supervisor 1)*

> *Upper management uses what we call a long screwdriver to turn every CO in the hierarchy in the same position, so to speak. This is how they micro-manage a situation deep into the capillaries of our organization. It's a lack of trust and a refusal to let things go and be dealt with at the appropriate level (Supervisor 7).*

On top of that, outside forces can sometimes press the chain of command to be retributive, for example, to satisfy public opinion:

> *There is a societal pressure to be professional and decisive. And that that decisiveness translates into "what happened to the culprit?" If you reply, "Well, nothing, but we learned a lot," no one buys that. (Focus group 1)*

Within the DoD, the RNLAF has a pyramid-shaped distribution of personnel meaning that as one progresses through the ranks, there are fewer equals in rank, culminating in just one person, the chief commander of the Air Force. Competition is fierce. Promotion is based on such formal aspects as years of experience, selection, and performance assessments. It is also based, partly, on informal aspects, such as the impression one makes on superiors. One respondent explained that COs are expected to deal with incidents in a decisive, manly, and task-oriented manner. Other respondents confirmed that restorative leadership is considered weak, inefficient, and time-consuming whereas retributive leadership is considered strong leadership.

> *We all saw what happened to that CO who was very outspoken [on social safety]. His career was derailed. (…) It nearly happened to me. (…) I did what I thought was right, but behind my back I was seen as too soft and that held up my career. (Supervisor 6)*

> *When something huge happens, you can feel and see convulsions, for the lack of a better word. The higher up the hierarchy you go, the bigger the convulsion. Basically, we're not allowed to do anything, every move we make is under a microscope. (Employee 1)*

The above quotes demonstrate how hard it can be for a CO trying to foster a restorative just culture when a retributive response is expected and desired.

For COs, withstanding the pressure to act according to retributive norms demands (moral) courage and the willingness to put the interest of the restorative just culture above personal career interests. This can demand an almost self-sacrificing attitude.

4.3 The tension between learning and punishing

The third theme relates to the tension between learning and punishing. The interviews revealed that the RNLAF considers the concept of just culture as vital for the organization, but it has not specifically stipulated the restorative approach to a just culture. Several respondents stated the importance of avoiding a punitive approach to ensure learning opportunities.

> *I think a just culture is an environment where people dare to be open, and dare to share their experiences, and everyone stays away from the question who is to blame. (…) The focus should be how are we going to learn from this, and which safety systems are we going to improve. (…) For me just culture is the opposite of blaming. (Supervisor 2)*

> *You must do everything you can to avoid having to punish someone (Supervisor 4).*

On top of that, sometimes it is not clear whether something is a punishment or not. The supervisor of pilot X explains that for many colleagues it was not clear whether pilot X was punished or not.

> *If we had repatriated him just to make him go home, then it would have been punishment indeed. But we really wanted to know: what chain of events led to this? What were your reasons? (…) That was something [we needed] to explain to the club, that it was not a punishment. The goal was to learn from this incident together. That remains a difficult tension because people still see it as punishment. (Supervisor 3)*

Respondents reported that sharing information, for example during a safety standdown, can feel like punishment.

> *Sharing your story can have the opposite effect. Standing in front of an audience during a safety standdown can be a form of punishment instead of an opportunity to learn from one another. (Focus group 4)*

As the interviews show, learning and punishing are seen as mutually exclusive. If one chooses to punish, the learning opportunity is lost. Not just from the individuals involved, but also from colleagues, because seeing people being punished discourages the intention to openly share information on occurrences. At the same time, sharing information about an occurrence can feel like a punishment because people are singled out to share their story, which again diminishes or even loses the learning opportunity.

4.4 A restorative response in a retributive context

The fourth theme relates to responding restoratively in a retributive context, which creates a hybrid form of just culture. Colleagues of pilot X said that past COs had tried to combine concerns for their own career with the just culture of the squadron. One respondent told how he was involved in a fly-by (a flight demonstration that often seeks to entertain or convey a thank you to the public on the ground) that the media judged as unnecessarily risky. Although he was cleared restoratively in his own squadron, after his CO had his annual review with his own superiors, his CO said to him:

> *"Okay, listen. Remember that incident last year? When the Commission of Inquiry comes, just bend over, and take one for the team." And I thought, Wait a minute. Am I being thrown under the bus here? (…) You notice that certain people have motives that are not in line with the collective well-being. (Focus group 1)*

The respondents describe this as cancelling the restorative progress made.

> *Focussing on your career and how your prospects can be impacted by the behaviour of the people under your command (…) really hampers the just culture in our organization (Supervisor 5).*

> *In the past we had different COs, whose first response was to see if they had to notify the military police. That definitely changes the climate in a squadron. Trust diminishes and it has a negative impact on the willingness to report. (Supervisor 2)*

Several respondents shared similar examples:

> *There was an incident. (…) We solved it in a certain way. (…) But there's a commander higher in hierarchy as well. (…) He just demanded that someone should be punished whereas I thought, you shouldn't do that, because it's counterproductive (Supervisor 4).*

> *[There was an incident] We had a morning briefing to discuss things there, but our supervisor told us: "Don't make an official report in the systems yet, because first we'll examine what happened ourselves. We don't want this to come out." So initially there was a bit of a cover-up. I thought that was quite strange. (Focus group 2)*

As the interviews show, responding restoratively inside a squadron, hidden from the rest of the world, but at the same time publicly responding retributively leads to diminished trust. The retributive response at least partly nullifies the effects of a restorative response.

5. Discussion

The results elicited four themes: 1) The need for vulnerability, both observed in the safety standdown and confirmed in the interviews, supports a restorative response but is hard to show in the retributive RNLAF climate; 2) the moral courage of leaders within the entire chain of command to act restoratively, withstanding the pressure to punish and the fear for one's own career when refraining from a task-oriented response; 3) the tension between learning and punishing which are considered mutually exclusive because punishment, or even just blaming, is detrimental to openness; and 4) a restorative response in a retributive context resulting in a hybrid form of a just culture which feels unsafe and confusing for the operators and results in diminished trust.

The RNLAF already has a strong awareness of the importance of just culture for (flight) safety. The concept is widely supported in the organization, which in recent years has invested in establishing this culture. For example, investigations focus on why the occurrence happened, acknowledging the fact that pilots and crew members are sincere professionals who can make mistakes and that the context of the occurrence should be considered. The RNLAF is also invested in the concept of speaking up to authority, both inside and outside the cockpit. Both developments fit the restorative just culture perspective. Still, the response to undesired outcomes is often retributive as the four themes emerging from our analysis demonstrate. We will provide an in-depth discussion of these themes in this chapter.

5.1 The need for and price of vulnerability

On the one hand, being vulnerable can lead to open discussions, repairing and learning. On the other hand, it can feel like a punishment because of the attention one draws and the resulting fear of retribution. Fear of retribution limits the free flow of information and thus inhibits organizational learning (Lawrenson & Braithwaite, 2018). This fear is confirmed when one is confronted by retributive elements in an intended restorative setting, such as being singled out in a tribunal or being required to answer blunt, judgemental questions from the audience. These retributive elements hamper the desired openness. A setting that expresses equality, for instance a circle that lacks specific positions (see Appendix 1), between all participants will support the restorative goal as will addressing the negative impact certain questions may have on the possibility to learn. Still, being vulnerable can be difficult. Besides having an encouraging setting, having a facilitator help prepare and guide the process can be useful, as is setting the ground rules (taking turns, no judging) to follow in this kind of gathering (see Appendix 1).

The difference in vulnerability seen in the two speakers relates to the concepts of apology and accountability. A narrative, vulnerable style shows remorse and taking responsibility for what happened as opposed to a distant and impersonal style. Based on a literature review, Prothero and Morse (2017) provided a list of components that must be present in an apology for it to heal damaged relationships. These include acknowledgement of an error, expression of sorrow, acceptance of responsibility by the

offender, validation of the impact on the people affected by the incident, an explanation of what happened, reason for the behaviour, sincerity, and a commitment to prevent recurrence (see also Appendix 1). All of these were present in the safety standdown. Regarding accountability, the literature shows that to give an account means providing reasons for shown behaviour (Messner, 2009). Accountability refers to the importance of the relationship between the people that exchange reasons for their behaviour (Johansen, 2008, Arbin et al., 2021).

5.2 Moral courage of leaders

The results illustrate how a hierarchical structure can impede an organization's ambition to foster a restorative just culture. The quotes show that COs feel they are trapped in a retributive climate. A retributive response is firm and decisive leadership, while a restorative response is soft and not associated with military leadership skills. Responding in a soft, restorative way poses the risk of derailing one's career. If COs want to impress their superiors by showing how they personally managed tricky situations, retribution looks better. At the same time, COs feel that when something negative happens, they can no longer be the leaders they aspire to be. Decisions on what should be done are made higher up the chain of command and the COs are tasked with executing the wishes of their superiors without having any say in the matter. Therefore, being able to respond restoratively requires either the support of the chain of command or moral courage and a self-sacrificing attitude in the CO. If the restorative just culture is not widespread practice, a CO will be resented for responding restoratively, as our data has demonstrated.

One challenge is dealing with residual retributive tendencies throughout the whole chain of command. Now, the way the RNLAF hierarchy works gets in the way of its attempts to foster a restorative just culture. The literature on restorative just culture shows that responding to occurrences requires asking the 'what' and 'why' questions in order to learn and restore (Dekker, 2016). This means that the focus is not just on an occurrence but also on the context in which it happened. Remarkably, the same restorative just culture literature does not seem to explicitly address the context of an organization that aspires to build a restorative just culture. The struggle an organization has when trying to adopt a restorative just culture in a predominantly retributive setting is not well discussed. The struggle results from the pressure to act retributively but also from being divided on the right course of action throughout the chain of command and amongst colleagues (Appendix 1). Real commitment from upper management is necessary to win that battle, not just in words but also in deeds. The belief in the benefits of restorative justice needs to be felt throughout the chain of command, or else restorative justice will be no more than the attempts of a few self-sacrificing leaders who have come to terms with ending their career because "it is the right thing to do for their staff." However, fostering a restorative just culture by sacrificing oneself seems hardly fair and a self-sacrificing attitude cannot be expected from every individual. The highest level of the chain of command can and should make the difference. As Beer explains (Beer, 2013), to change an organization, senior management starts with committing itself to honest conversations about organizational strengths as well as weaknesses that cause underperformance. Because of the perceived unjustness and subsequent hiding of information and lack of learning, a retributive climate will lead to underperformance. Fostering a restorative just culture must then start with senior management dealing productively with these challenges.

5.3 The tension between learning and punishing

As mentioned in the introduction, according to Pupulidy the currency of safety is information. If that currency is shut down, if the information flow stops because of the fear of retribution, learning and hence safety will be negatively impacted. A tension between learning and punishing was observed during the safety standdown and found in the interviews. As articulated in the literature (Heraghty et al., 2020), the respondents perceive these two concepts as mutually exclusive. Punishment emphasizes the unreliability of humans, the deviance of failures and the uniqueness of the offender. It motivates people to not get caught next time. It restricts safety lessons to a specific part of the organization, and it is about closure, leaving the unwanted event in the past. Dekker also mentions less willingness to report incidents if employees think that their reports are treated unfairly or lead to negative consequences (Dekker, 2009). Following the safety standdown, the employees' comments on not speaking up at the next safety standdown is an example of this.

In contrast, learning highlights the safety that humans create. It focusses on the fact that failures are normal in the complex sociotechnical system to which every operator is exposed. Learning generalizes safety lessons for the entire organization. The accent lies on avoiding a future event instead of avoiding getting caught. It is about continuous improvement, not leaving the unwanted event in the past but incorporating it in future practices (Dekker, 2003). The findings show that sharing experiences and deliberations fosters an understanding of the role of a complex system. To achieve that, the right kind of questions need to be asked. Dekker (2016) observed that a retributive just culture focuses on the individual (who) and not on the situation (what and why), which is common in a restorative just culture. This spontaneous shift in focus of the questions in the first presentation, from 'who' to 'what' and 'why' in the narrative of the second speaker and the acknowledging tone of the subsequent questions made it possible to learn from the experience of a colleague. To ensure learning occurs during subsequent safety standdowns it would be helpful to change the underlying question of individual presentations from "what have I experienced/done?" to "what did you encounter that could happen to any one of us?" Actively expressing this second question would emphasize the intention of learning.

5.4 A restorative response in a retributive context

Creating a restorative just culture in an organization that has a retributive tradition is challenging, as this research shows. One result of this complex process could be a form that mixes punishing and learning and misses the opportunity to learn. In other words, this hybrid threatens the development of a restorative just culture, as observed in the RNLAF. While it stipulates a focus on learning, on looking forward and avoiding blame, therefore propagating a restorative just culture, when an incident occurs the response, whether formal or informal, is mostly retributive, ranging from gossip through judgement to actual punishment.

When it comes to occurrences that pertain to criminal law, every organization in the Netherlands has no choice in how to respond. In these instances, the prosecution decides if a behaviour is punishable. Luckily, this does not happen often in the RNLAF. Most incidents fall within the realm of disciplinary or administrative law, so that the RNLAF can decide how to solve the matter. When this is the case, it is most certainly possible to respond restoratively but this choice is not often made.

Surprisingly, the promising results from the realm of criminal law show that restorative justice is increasingly sought as a means of dealing with a crime, bringing together the offender, victims, relevant support people and facilitators (Zinsstag et al., 2011). Here restorative justice is seen as "a normative response to actions that have caused significant harm to individuals, relationships, and the community" (Ward & Heffernan, 2017, p.46). Examples are Family Group Conferencing in the case of a young offender (Carswell et al., 2013) or in cases of child abuse, neglect and offences in Aotearoa/New Zealand (Becroft, 2017), sentencing circles involving members of a community, family members, victims, and a judge (Laprairie, 1995) and a 'balanced model' in youth justice in Northern Ireland (Zinsstag & Chapman, 2012).

If restorative justice is deemed possible by criminal law, it is certainly possible under disciplinary or administrative law. What inhibits this development is the idea that punishment is incompatible with restorative just culture. However, this is not the case. The remarkable thing about restorative just culture is that the process can still end in measures being taken regarding the people involved. The difference with the retributive or hybrid forms is that the people involved in the occurrence are also involved in determining what the consequences should be and whether they are appropriate. After hearing and acknowledging the accounts of everyone affected, measures are based on this profound understanding of the situation and feel 'just' to the people involved. Acknowledging the true extent of restorative just culture, and therefore preventing a hybrid form, will help in overcoming the complexities faced when trying to foster one.

If it can be such a struggle to create a restorative just culture in a retributive context, why should an organization still aim for restorative just culture? Because of the long-lasting effects of learning and restoring, as observed in this research: pilot X is once again a respected member of his community. Pilot X acknowledged his role in the occurrence and expressed his sincere remorse, both of which made him feel very uncomfortable. This is how he faced 'retribution' in a restorative climate (see Appendix 1). The results seen in this case study are substantiated by the literature regarding the effects of restorative justice. The aim of restorative just culture is to repair and learn (Dekker, 2016; Dekker & Breakey, 2016). Besides the positive impact regarding trust (Turner et al., 2020), available information (Dekker & Breakey, 2016), safety (Bell et al., 2011), suspensions and disciplinary measures, absence, retention and cost savings (Latimer et al., 2005) and reduced youth offences (Mulligan, 2009), Kaur et al. (2019) and Stang et al. (2013) found higher compliance rates, a reduction in recidivism and improved satisfaction with the outcome as a result of restorative justice for all people affected. Ward et al. (2014) explain the reduced recidivism (Shapland et al., 2008; Strang et al., 2013) as a possible result of a shared sense of norms and values between offender and the people affected as well as the associated actions needed to repair the damage done and restore the relationships that were hurt. Nugent and Schinkel (2016) report that if others recognize a change in the offender it has an enormous positive impact. Kirkwood (2021) explains that the relational process of restorative justice provides the opportunity for an offender to explain to their community how the moral standards of this community will have an impact on their future behaviour.

The observations, the interviews and the literature study show that restorative justice not only has immediate positive consequences for a specific situation, but it also has long-term positive consequences in terms of more openness and subsequently more safety. Appendix 1 contains a list of transferable factors, based on the safety standdown, which can help make an intervention as restorative as possible.

6. Strengths and limitations of the study

In this paper we presented the results of a case study. The strength of this method is that it gives a rich – detailed, in depth – description of a certain situation. Another value is that this intensive type of qualitative research gives insight into the lived experience of people, and often sheds new light on how existing theories work out in practice. In our study we unravelled the complexity of remaining consistently restorative in a culture that is traditionally retributive. A limitation of this study is that it pertains to only two presentations. Although the difference in style and the instant, spontaneous change in approach (from anger and distrust to understanding and forgiveness) were remarkable, it is possible that the difference in audience response in this safety standdown was influenced by other factors than the vulnerability expressed, for instance the presenter's personality or position within the group. We tried to counter this limitation by conducting additional interviews. What makes our data robust is that we interviewed not just one person, but various people involved in the safety standdown: the culprit, his colleagues, his supervisor. The relatively large number of interviews, for qualitative standards, ensured that we could analyse the safety standdown from different perspectives, which reduced bias. We conducted 13 interviews in all, involving 29 people in total: interviews specifically focussed on the safety standdown (n=4), individual interviews (n=5) and focus group interviews on just culture in the RNLAF in general (n=4).

Another limitation of this study is that it pertains to a very local situation – a safety standdown of a squadron of the RNLAF –, and that it is unclear what the generalisability is. We describe a safety standdown of a squadron of the RNLAF. Further research has to enlighten to what extent the results are generalizable to other branches of the Netherlands Armed Forces, or to military organizations in other countries. However, this case study was part of a larger study involving units in other branches of the Dutch Armed Forces. Many of the themes presented here – the value of vulnerability, the importance of courageous leaders, resisting the tendency to punish – where also mentioned by other respondents in other units within the Dutch Armed Forces.

Collecting more case studies will give more insight into the complexities surrounding fostering a restorative just culture in dominantly retributive settings. In the broader action research project aimed at fostering a restorative just culture in the Dutch Department of Defence we are already collecting best practices as learning opportunities. The lessons learned from our case study of the complexities surrounding implementing a restorative culture could be relevant for other hierarchical, competitive settings that are traditionally retributive. However, further research is needed.

7. Conclusion

The complexities of fostering a restorative just culture in practice were examined in this case study of a safety standdown in the RNLAF. The findings show that four factors are of great importance for a restorative response: 1) a willingness to be vulnerable without being punished for this; 2) moral courage and a fair attitude, even if it means sacrificing one's own interests; 3) a management that does not

punish COs for responding restoratively; and 4) a desire to learn from the mistake instead of punishing for it. The literature suggests that restorative response has positive effects in both the short and long term. This was also visible in this case study: relationships and trust were restored; the reasons for the decisions made were understood; the community learned from the incidents; and gossip and calls for punishment ceased.

A restorative response – thus a restorative just culture – is aimed at attaining a greater safety level in the RNLAF. However, this study demonstrates how hard it is to aspire to a restorative just culture in a predominantly retributive setting and culture. Examples that fostered restorative openness were threatened by the deeply engrained response to punish in the RNLAF. This makes this safety standdown even more a rare and exceptional example. Although it was intended to create a restorative setting, it only partially accomplished this goal and even then, only spontaneously [by chance]. Therefore, this particular safety standdown cannot be seen as a restorative measure but as a restorative response within an overall retributive setting. This analysis shows the complexities the RNLAF, and other organizations, can learn from and need to overcome to strengthen a restorative response and build a restorative just culture that will enhance the safety of the organization.

References

Alicke, M. D., Buckingham, J., Zell, E., & Davis, T. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin*, *34*(10), 1371–1381. https://doi.org/10.1177/0146167208321594

Arbin, K., Frostenson, M., Helin, S., & Borglund, T. (2021). Explaining workers' resistance against a health and safety programme: An understanding based on hierarchical and social accountability. *Safety Science*, *136*(March 2020), 105131. https://doi.org/10.1016/j.ssci.2020.105131

Balcerzak, T. (2017). *Scientific Journal of Silesian University of Technology . Series Transport DOES ELECTRIC CAR PRODUCE EMISSIONS ? 94*.

Becroft, A. (2017). *Family group conferences: Still New Zealand's gift to the world?* http://www.occ.org.nz/assets/Uploads/OCC-SOC-Dec-2017-Companion-Piece.pdf

Beer, M. (2013). The Strategic Fitness Process: A Collaborative Action Research Method for Developing and Understanding Organizational Prototypes and Dynamic Capabilities. *Journal of Organization Design*, *2*(1), 27. https://doi.org/10.7146/jod.8017

Bell, S. K., Delbanco, T., Anderson-Shaw, L., McDonald, T. B., & Gallagher, T. H. (2011). Accountability for medical error: Moving beyond blame to advocacy. *Chest*, *140*(2), 519–526. https://doi.org/10.1378/chest.10-2533

Braithwaite, J. (2002). *Restorative Justice and Responsive Regulation*. Oxford University Press.

Brborović, O., Brborović, H., Nola, I. A., & Milošević, M. (2019). Culture of blame—an ongoing burden for doctors and patient safety. *International Journal of Environmental Research and Public Health*, *16*(23). https://doi.org/10.3390/ijerph16234826

Carswell, S. L., O-Hinerangi, M., Gray, C., & Taylor, A. (2013). *Evaluation of Family Group Conference Practice & Outcomes - Scoping phase*. https://thehub.swa.govt.nz/assets/documents/42676_evaluation-of-fgcs-scoping-study-25-9-14_0.pdf

Catino, M., & Patriotta, G. (2013). Learning from Errors: Cognition, Emotions and Safety Culture in the Italian Air Force. *Organization Studies*, *34*(4), 437–467. https://doi.org/10.1177/0170840612467156

Charmaz, K. (2011). A Constructivist Grounded Theory Analysis of Losing and Regaining a Valued Self. In *Five Ways of Doing Qualitative Analysis: Phenomenological Psychology, Grounded Theory, Discourse Analysis, Narrative Research, and Intuitive Inquirey.* The Guilford Press.

Dekker, S. W. A. (2003). When human error becomes a crime. *Human Factors and Aerospace Safety*, *3*(1), 83–92.

Dekker, S. W. A. (2006). *The Field Guide to Understanding Human Error*. Ashgate Publishing Limited.

Dekker, S. W. A. (2007). *Just Culture; Balancing Safety and Accountability*. Ashgate Publishing

Limited.

Dekker, S. W. A. (2009). Just culture: Who gets to draw the line? *Cognition, Technology and Work*, *11*(3), 177–185. https://doi.org/10.1007/s10111-008-0110-7

Dekker, S. W. A. (2016). *Just culture: Restoring Trust and Accountability*. CRC Press Taylor & Francis Group.

Dekker, S. W. A., & Breakey, H. (2016). "Just culture:" Improving safety by achieving substantive, procedural and restorative justice. *Safety Science*, *85*, 187–193. https://doi.org/10.1016/j.ssci.2016.01.018

EUROCONTROL. (2006). Establishment of "Just Culture" Principles in ATM Safety Data Reporting and Assessment - EAM2/GUI 6. In *Establishment of "Just Culture" Principles in ATM Safety Data Reporting and Assessment - EAM2/GUI 6.* (Issue March 2006). https://skybrary.aero/sites/default/files/bookshelf/235.pdf

Goodstein, J., & Aquino, K. (2010). And restorative justice for all: Redemption, forgiveness, and reintegration in organizations. *Journal of Organizational Behavior*, *31*(4), 624–628. https://doi.org/10.1002/job.632

Heraghty, D., Dekker, S. W. A., & Rae, A. (2021). Modifying an accident process and its justice system – From single narratives and retribution to multiple stories and restoration. *Safety Science*, *139*(April 2020), 105248. https://doi.org/10.1016/j.ssci.2021.105248

Heraghty, D., Rae, A. J., & Dekker, S. W. A. (2020). Managing accidents using retributive justice mechanisms: When the just culture policy gets done to you. *Safety Science*, *126*(May 2019), 104677. https://doi.org/10.1016/j.ssci.2020.104677

International Civil Aviation Organization. (2018). *FOSTERING JUST CULTURE IN OPERATORS AND SERVICE PROVIDERS*. https://www.icao.int/APAC/Meetings/2018 APRAST12/APRAST12 WP-11 AI_5 - [NokScoot_AEROTHAI_CANSO] Just Culture.pdf#search=just culture

Johansen, T. R. (2008). Employees and the operation of accountability. *Journal of Business Ethics*, *83*(2), 247–263. https://doi.org/10.1007/s10551-007-9615-x

Kaur, M., De Boer, R. J., Oates, A., Rafferty, J., & Dekker, S. W. A. (2019). Restorative Just Culture: a Study of the Practical and Economic Effects of Implementing Restorative Justice in an NHS Trust. *MATEC Web of Conferences*, *273*(April 2018), 01007. https://doi.org/10.1051/matecconf/201927301007

Kirkwood, S. (2021). A practice framework for restorative justice. *Aggression and Violent Behavior*, *November*, 101688. https://doi.org/10.1016/j.avb.2021.101688

Kováčová, M., Licu, A., & Bálint, J. (2019). Just Culture - Eleven Steps Implementation Methodology for organisations in civil aviation - "jC 11." *Transportation Research Procedia*, *43*, 104–112. https://doi.org/10.1016/j.trpro.2019.12.024

Laprairie, C. (1995). Altering course: New directions in criminal justice: Sentencing circles and family

group conferences. *Australian & New Zealand Journal of Criminology*, *28*.

Latimer, J., Dowden, C., & Muise, D. (2005). The effectiveness of restorative justice practices: A meta-analysis. *The Prison Journal*, *85*(2), 127–144. https://doi.org/10.1177/0032885505276969

Lawrenson, A. J., & Braithwaite, G. R. (2018). Regulation or criminalisation: What determines legal standards of safety culture in commercial aviation? *Safety Science*, *102*(July 2016), 251–262. https://doi.org/10.1016/j.ssci.2017.09.024

Lipira, L. E., & Gallagher, T. H. (2014). Disclosure of adverse events and errors in surgical care: Challenges and strategies for improvement. *World Journal of Surgery*, *38*(7), 1614–1621. https://doi.org/10.1007/s00268-014-2564-5

Marx, D. (1997). Discipline: The role of rule violations. *Ground Effects: Reporting Maintenance and Groundcrew Error Reduction Efforts*, *2*, 1–3. http://www.system-safety.com/articles/GroundEffects/Volume 2 Issue 4.pdf

Messner, M. (2009). The limits of accountability. *Accounting, Organizations and Society*, *34*(8), 918–938. https://doi.org/10.1016/j.aos.2009.07.003

Mulligan, S. (2009). From Retribution to Repair: Juvenile Justice and the History of Restorative Justice. *University of La Verne Law Review*, *31*(1), 139–149.

Nugent, B., & Schinkel, M. (2016). The pains of desistance. *Criminology and Criminal Justice*, *16*(5), 568–584. https://doi.org/10.1177/1748895816634812

Parkinson, J., & Roche, D. (2004). Restorative justice: Deliberative democracy in action? *Australian Journal of Political Science*, *39*(3), 505–518. https://doi.org/10.1080/103614042000295101

Prothero, M. M., & Morse, J. M. (2017). Eliciting the functional processes of apologizing for errors in health care: Developing an explanatory model of apology. *Global Qualitative Nursing Research*, *4*. https://doi.org/10.1177/2333393617696686

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality and Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Shapland, J., Atkinson, A., Atkinson, H., Dignan, J., Edwards, L., Hibbert, J., Howes, M., Johnstone, J., Robinson, G., & Sorsby, A. (2008). *Does restorative justice affect reconviction? The fourth report from the evaluation of three schemes* (Issue June). www.justice.gov.uk/publications/research.htm

Strang, H., Sherman, L. W., Mayo-Wilson, E., Woods, D., & Ariel, B. (2013). Restorative Justice Conferencing (RJC) Using Face-to-Face Meetings of Offenders and Victims: Effects on Offender Recidivism and Victim Satisfaction. A Systematic Review. *Campbell Systematic Reviews*, *9*(1), 1–59. https://doi.org/10.4073/csr.2013.12

Turner, K., Stapelberg, N. J. C., Sveticic, J., & Dekker, S. W. A. (2020). Inconvenient truths in suicide

prevention: Why a Restorative Just Culture should be implemented alongside a Zero Suicide Framework. *Australian and New Zealand Journal of Psychiatry*, *54*(6), 571–581. https://doi.org/10.1177/0004867420918659

Ward, T., Fox, K. J., & Garber, M. (2014). Restorative justice, offender rehabilitation and desistance. *Restorative Justice*, *2*(1), 24–42. https://doi.org/10.5235/20504721.2.1.24

Ward, T., & Heffernan, R. (2017). The role of values in forensic and correctional rehabilitation. *Aggression and Violent Behavior*, *37*, 42–51. https://doi.org/10.1016/j.avb.2017.09.002

Woods, D. D., Dekker, S. W. A., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Ashgate Publishing Limited.

Zinsstag, E., & Chapman, T. (2012). Conferencing in Northern Ireland: Implementing restorative justice at the core of the criminal justice system. In E. Zinsstag & I. Vanfraechem (Eds.), *Conferencing and restorative justice: International practices and perspectives.* (pp. 173–188). Oxford: Oxford University Press.

Zinsstag, E., Teunkens, M., & Pali, B. (2011). *Conferencing: A way forward for restorative justice in Europe.* https://www.euforumrj.org/sites/default/files/2019-11/final_report_conferencing_revised_version_june_2012.pdf

Appendix 1: Transferable factors: establishing a restorative reaction to occurrences

Reflecting on the safety standdown as an example of the challenges and success in fostering a restorative response, the following transferable factors can be distinguished:

When the incident occurs

- Keep the time between an occurrence and a collective learning opportunity, such as a safety standdown, as short as possible to reduce gossip and informal judgements.
- During a safety standdown, the entire organization, including the audience, should support a restorative view. A call from colleagues for a retributive response derails the restorative just culture.
- Recognize the pressure coming from stakeholders outside the community, the overall organizational culture (competitive, hierarchical) and address their influence on the restorative process one is trying to achieve.
- A restorative response is a team effort: upper management and middle management need to be on the same page. A retributive call from superiors can derail the restorative process. Make sure the restorative approach is supported throughout the chain of command according to the provided mandate.
- Use the concepts of ownership and moral courage to diminish the impact of the occurrence, when using the mandate given to deal with the situation as one sees fit.

The adage the squadron management lives by contributes to the safety standdown process. Management agrees to react in a certain way if the unexpected, usually undesired, happens:

- They will calmly note that something has happened but fetch a cup of coffee first;
- They will not react impulsively but apply the concept of *audi alteram partem* (fair hearing) to gain insight into a situation;
- They will spend time deliberating before setting a course. They are willing to change this course or even turn back from it if the results don't live up to expectations.

This agreement can be seen as a transferable factor. Reacting like this makes a restorative response not just possible but likely.

Preparing the gathering and the speaker

- It can be useful to ask someone from outside the community to facilitate the restorative process and ensure that it does not turn into a retribution panel. An outside facilitator has the advantage of not feeling the pressure to act in a certain way, because this person is not part of the particular chain of command (as is leadership for instance) or part of the group process. However, given that the facilitator is part of the RNLAF, inside knowledge and understanding of context and operations is nonetheless present.

- Create a socially safe restorative setting that allows each participant to share his thoughts In addition to having a skilled facilitator and setting some ground rules (next section), choose an appropriate room and seating, e.g., form a circle instead of a theatre (stage/audience) set-up.
- The importance of the leader/superior during the restorative event is no different than that of any other participant. Restorative justice is not about rank or position but about hurt, what is needed and forward-looking. Make sure that the leader is not awarded a specific position in either seating or conversation. Nor should the 'culprit' be placed centre stage.
- Use intimate narratives to tell the story, without PowerPoint or a lectern, to connect with the audience. If necessary, train the speaker in this.
- For a restorative response, important aspects are authenticity, a sincere explanation and apology and acknowledgement that others may have been affected by one's actions. Disclosing personal reasons, feelings and experiences as facts is also important.
- To prepare the speaker, one can use Porthero and Morse's list of components that an apology needs to possess in order to heal damaged relationships (Prothero & Morse, 2017). These are: acknowledgement of an error, expression of sorrow, acceptance of responsibility by the offender, validation of the impact on the people affected by the incident, an explanation of what happened, reason for the behaviour, sincerity, and a commitment to prevent recurrence. Furthermore, the facilitator can support the speakers by using their knowledge of restorative justice and explicitly expressing support and understanding for the occurrence about to be shared.

Setting the rules of the gathering

- Discuss the ground rules for sharing information and thoughts during the safety standdown or similar event, e.g., respectfully taking turns, no interrupting, and no judging.
- Explicitly mention the learning and healing goal of a safety standdown or similar event.
- Explicitly mention all the people affected by the occurrence, whether they be first or second victims.
- Be open and honest about how one is affected. For instance, a commanding officer can fear the consequences of the way they are dealing with the occurrence for their own career.
- Encourage all participants to take on the perspectives of all other participants

During the gathering

- The 'culprit' needs to be supported in expressing his vulnerability.
- Allow feelings of discomfort. Healing does hurt. This means that in the restorative response 'retribution' lies in acknowledging, without reticence, one's own role in the occurrence. As seen in the case study this vulnerable and uncomfortable position is necessary for healing to occur between colleagues.

General

- Restorative leadership should gain one higher status. Historically, there are many inspiring examples. Share these to ensure that restorative leadership is rewarded in the chain of command.

Chapter 6. Reducing gaps between paper and practice requires more than a technical alignment

Leonie Boskeljon-Horst, Robert Jan de Boer, Sidney Dekker (n.d)

Safety Science (under review)

Reducing gaps between paper and practice requires more than a technical alignment

Abstract

The way work is done often differs from the way work is imagined. An enduring difference between practice and prescription is considered undesirable, from both the organisational and academic point of view. In this study we aimed to reduce the gap between work-as-done (practice) and work-as-imagined (paper) by utilizing a technical alignment approach that combines micro-experiments with a framework for rule management that places monitoring and adapting rules key to the process. In four cases, we tested whether technical alignment could reduce the gap, change the rules and break through organisational complicity. We found that the micro-experiments were technically successful as they revealed the effectiveness of interventions, generated shop floor commitment and reduced gaps. However, we also found that this approach overlooked the difficulty of actually changing rules and organisational complicity remained. Adding insights obtained from a fifth case discovered during the research, we conclude that technical alignment should be supplemented by a focus on improving psychological safety, implementing a restorative just culture and upholding psychological contracts. Lasting change requires leaders to display moral courage throughout the entire chain of command.

1. Introduction

When trying to understand human work it is useful to differentiate between the way work is actually performed, also called work-as-done (Moppett & Shorrock, 2018; Shorrock, 2016) and how it should be done according to the way it is prescribed, often in rules and procedures, also called work-as-imagined (Hollnagel, 2012; 2014). Although rules and procedures are considered important – they help people to remember steps in a task under varying circumstances, ensure that people can work together effectively, function as organisational memory to indicate how processes work and as a means of identifying deviations that might be problematic (Hale & Borys, 2013a) – they cannot capture the dynamics and complexities of every work situation (Bieder & Bourrier, 2013; Damoiseaux-Volman et al., 2021b).

Work takes place in a context of limited resources, multiple goals and pressures as well as surprises, constraints and novelty (Loukopoulos et al., 2009). In many cases, it can be impossible to follow the rules and get the job done at the same time (Bieder & Bourrier, 2013; Dekker, 2003a). For instance, Tucker (2004) describes how disruption and errors in information hamper employees in doing their job. These 'operational failures' result in necessary rather than voluntary workarounds (see also (Halbesleben et al., 2008; Tucker et al., 2020)).

As a result of contextual variation, the way work is done often differs from the way work is imagined and prescribed (Damoiseaux-Volman et al., 2021b; Hollnagel, 2014). Employees come up with creative

solutions to get the job done in spite of the rules and procedures. Creative solutions are often called violations (Reason, 1997), deviations (Edgett Collins, 2007), short-cuts (Gauthereau & Hollnagel, 2005) or workarounds (E. W. Morrison, 2006). Each term stresses the undesirability of work-as-done differing from work-as-imagined.

Although working differently in practice than prescribed can actually be good in some situations (e.g. Besnard & Greathead, 2003), an enduring difference between practice and prescription as a result of hampering rules is considered undesirable from an organisational point of view. This is because in these cases ill-functioning rules and process innovations are actively hidden from management and bypass the controls for risk inherent in the rules. A section of the scientific safety community has the same concern regarding the undesirability of enduring gaps between prescription and practice (Hale & Borys, 2013a).

In a previous study at an operational squadron of a military organisation, we found conflicting goals resulted in local ingenuity to get the job done (Boskeljon-Horst et al., 2022). These enduring gaps (non-normatively labelled 'local ingenuity' by the authors) improved goal realisation and were consistent between operators. Neither management nor operators made any active attempt to reduce the gaps between prescription and practice even if the organisation considered them undesirable. One explanation for the persistence of these gaps is that while operators utilise an adaptive paradigm towards rules to cope with the complexities of work, management actively disregards this reality and instead abides by a classical paradigm focused on compliance – possibly out of political necessity. In another previous study the authors found that the political focus on compliance seems reminiscent of a bogeyman story: 'good behaviour' is ensured by the threat (fear) of punishment (Boskeljon-Horst, De Boer, et al., 2023). Cases of noncompliance are best left undisclosed, and thus the need to modify rules remains obscured. Hale and Borys (2013a) use "organisational complicity" to define this self-perpetuating cycle of deviation and conspiratorial silence about poor rules.

1.1 Reducing the gap

In a follow-up paper Hale & Borys (2013b) present an approach that they suggest to be effective in challenging the cultural assumptions of top managers, supervisors and regulators and therefore reconciles the adaptive, traditional paradigms between managers and operators. Their framework for rule management focuses on monitoring rules and collaboratively adapting these as a means to close the gap between prescription and practice (in contrast to a focus on enforcing discipline in the use of rules). This framework is intended to combine the classical and constructivist paradigms of rule management. Its intention is to improve rule use, ensure rules match practice and foster a dialogue among rule makers, rule enforcers and rule users. The framework consists of nine steps and shows a smaller and a larger loop. In the smaller loop, rules are evaluated for their effectiveness, the use of good rules is enforced and exceptions are dealt with and continuously monitored. In the larger loop, bad rules are redesigned or scrapped. Based on risk scenarios, new rules developed, written, tested and approved and employees are trained in their use and adaptation. Although the focus is on rule monitoring (for matching with practice), adaptation and modification. Although the focus is on monitoring, adaptation and modification, the rule management framework does not provide certainty regarding the effectiveness of interventions focusing on adaptation and modification nor does it explicitly state how a structural dialogue among rule makers, rule enforcers and rule users is ensured. The current authors

have found no evidence in the literature of successful application of this approach in organisations suffering from organisational complicity.

Using micro-experiments is another approach that seems promising to reduce differences between work-as-done and work-as-imagined. As Snowden and Boone (2007) state, in a complex situation it is very difficult to infer the right solution upfront. Additionally, due to the experimental and bottom-up nature of micro-experiments, we speculate that this will facilitate the necessary conversations to encourage an adaptation of rules to working practice. Dekker (2018a) suggests that with 'safe to fail' experiments and carefully monitoring the results, the right course of action will reveal itself through an intuitive, iterative approach. These trial interventions have become known as micro-experiments, first used by Dekker in a project to enhance safety in Woolworth supermarkets as a means to test the elimination of safety controls (Dekker, 2018). Micro-experiments are a safe way to test interventions aimed at improving the working situation (Plioutsias et al., 2020), using the organisation's own workplace and employees (Dekker, 2018). They are different to ad hoc interventions or long term projects because they are planned before they are executed, and they are carefully monitored to see whether the intended results are realised. They are also different from large scale projects in that those are not supposed to fail, whereas micro-experiments are temporary and allow for unexpected, even disappointing results (De Boer, 2021). Micro-experiments are usually generated bottom-up because the users' context is an important factor (Christensen et al., 2016). Successful interventions are based on what operators believe they need to get the job done effectively and safely (Dekker, 2014). Micro-experiments are useful for an organisation to build confidence. They can show that a different approach to rule management is possible and, based on the gathered data, this can effectively reduce the gap between paper and practice (Dekker, 2018). Other than the referenced sources, further literature on effectively reducing the differences between work-as-done and work-as-imagined is limited.

1.2 Proposed framework

In the present study we investigate the effectiveness of combining the procedural, intuitive and iterative approaches to reduce the gap between the way work is done and how it is imagined. We adopt the rule modification methods described by Hale and Borys (2013b) and and combine this with micro-experiments (Dekker, 2018a) as proposed by De Boer (2021). This approach holds monitoring and adaptation central to the management of rules (rather than focusing on rule-use discipline) and uses micro-experiments to test improvements to rules. Although the framework proposed by Hale and Borys (2013b) describes how rules can be changed, we have found no evidence for its effectiveness. In particular, we have found that a Bogeyman belief will inhibit conversations about rule misalignment necessary to induce an adaptation of rules to working practice (Boskeljon-Horst, De Boer, et al., 2023).

As De Boer (2021) suggests, a different approach is required when the aim is to reduce a significant gap between work-as-imagined and work-as-done. Micro-experiments provide a means to test if interventions bring about the envisaged results. Originally, they were not intended to reduce gaps between work-as-imagined and work-as-sone, but to experiment with fewer safety controls (Dekker, 2018). The suggestion is that in using micro-experiments, it is possible to test new rules without the normative assessment of non-compliance and therefore get buy-in from stakeholders for a change to the rules.

The approach applied in the current study is novel in combining the constituent parts in an effort to reduce the gap between prescription and practice, change the rules to better match daily operations, and break through the organisational complicity evident in the target organisation.

1.3 Target organisation

This study was conducted in the same flight squadron of the Royal Netherlands Air Force (RNLAF) as employed in two previous studies. The first study examined the effect of two conflicting paradigms regarding rule perception and management on goal realisation in an operational setting through the identification of local ingenuity (Boskeljon-Horst et al., 2022). The second study examined why a large gap between the way work is done and the way it is imagined and prescribed endures despite its undesirability (Boskeljon-Horst, De Boer, et al., 2023).

The flight squadron is an operational unit containing 55 male employees, aged between 22 and 45, and some 30 aircraft. Squadron activities include flight planning, briefing, debriefing, line maintenance, and ancillary support. The squadron's home base is one of the Dutch airfields, but it is also active in other North Atlantic Treaty Organization (NATO) countries for joint exercises, and on deployment missions elsewhere. The goals of the squadron are to maximise the individual's potential by developing their skills and expertise, executing the missions assigned by the air force commander and his delegates within the appropriate budget, and performing them safely. An overall RNLAF goal is to follow existing regulations, rules and procedures.

The RNLAF has to deal with laws ranging from strict to less strict, such as the Aviation Act, and with civil regulations from the International Civil Aviation Organisation, the Joint Aviation Authorities and the European Aviation Safety Agency. The RNLAF is also impacted by the rules of the European Defence Agency, NATO standardisation agreements (STANAG) and the Military Aviation Authority, such as aviation regulations, acceptable means of compliance and guidance materials. Finally, there are RNLAF rules for aircraft and flight operations as well as exceptions or temporary rules, standard operating procedures, guidelines and checklists.

The RNLAF belongs to the Dutch Department of Defence (DoD), a political entity that has suffered from a decade of austerity, while political ambition and expectations of missions to be executed have remained the same. This has created tension within the DoD: budget cuts led to downsizing – which led to reduced inflow and many vacancies – and increasing pressure to deliver the expected output with a reduced work force.

1.4 Research aim

This research applies the framework proposed by Hale and Borys (2013b) in combination with micro-experiments. We seek to understand whether this method would allow us to reduce the gap between prescription and practice, change the rules to better match daily operations, and break through the organisational complicity evident in the target organisation.

Our study adds to both the literature and practice of of aligning rules with working practices. It compliments earlier theoretical and empirical work on rule violations by testing a novel approach in practice, with a focus on the safety domain. The results of this study will also inform practitioners who are charged with the alignment of rules and working practices.

2. Method

This investigation took the form of a multi-case study (Hak & Dul, 2007) in which multiple cases were analysed and common factors were identified to answer the research question. One internal researcher (researcher 1) and two external researchers (researchers 2 and 3) conducted the study. Researcher 1 knows the target organisation, customs and language, which facilitated understanding of the problem and goal. Researchers 2 and 3 have experience in conducting micro-experiments and provided an antidote to a possible blind spot of the internal researcher. The authors generally took a passive role in the execution of the micro-experiments and the modification of rules, and had access to meetings and meeting notes. In some cases the researchers' help was requested in filling out the micro-experiment template.

2.1 Selection of cases

The operational unit which is subject of the current investigation has been nominated by the military organisation for its creativity and proactiveness. The unit is battle-ready and has challenging operational, budgetary and compliance objectives. It is in good standing with the military leaders as an effective and socially-coherent unit, although it is also suspected that the unit needs to cut corners to successfully achieve all its goals (Boskeljon-Horst et al., 2022). The cases for the current investigation are selected based on the following criteria:

- Goal conflicts evident at the time of the study. To be identified in preliminary interviews with unit leadership.
- Specific enough for further analysis and description, to be ascertained in interviews with the front-line staff directly involved in the operation.
- Regarding real or possible gaps between work-as-done and work-as-imagined. In some cases, a violation with governing rules may be evident. In other cases, the goal conflicts are possibly solved by abiding by the rules, even though this is less efficient, effective or even less safe. Both categories are included in the selection.
- Enduring nature: we aim to ascertain during the interviews with the front-line staff that the issue has persisted at least over the last six months to ensure that we do not include temporary snags.

- We include an explicit desire of the target organisation's management to reduce the gap between work-as-done and work-as-imagined as criterion. To be determined in the initial interview with military leaders as part of the current study.
- The availability of employees to conduct the micro-experiments: To be determined in the initial interview with operators as part of the current study.

We trust that the resulting selection of cases is sufficiently numerous and divers to find common factors to answer the research question yet enable generic conclusions.

2.2 Micro-experiment execution

For each of the selected cases, a micro-experiment will be executed to test possible rule changes before a modification for the rules is formally proposed. The micro-experiment execution follows the approach proposed by De Boer (2021, p. 52). Each intervention begins with a startup meeting to define the problem to be tackled and identify multiple solutions. The intervention with the most promising cost-benefit ratio is chosen. Notes of each start-up meeting are made according to a template to capture this and to prevent deviation from the intended course of the intervention. The template helps to ensure the micro-experiment is kept manageably small, that constraints are in place to limit new risks, decide on success criteria, identify how to stop or enlarge the intervention, and set the end date. A baseline is identified to compare results with.

The results of each micro-experiment are used to decide on subsequent steps: formalise, modify or abandon the intervention. Data is collected by attending the start-up meetings, observing the interventions, supplemented by open-ended and in-depth interviews on the effectiveness of the interventions.

2.3 Rule adaptation

In case the micro-experiment leads to a proposal to formalise the rule change, the researchers will monitor the proposal to scrap, modify or add rules, as well as the discussions with rule makers and leaders that are aimed at getting approval for the implementation. Data is collected by open-ended and in depth interviews with front line staff, rule makers and leaders, and includes document analysis of the proposals. If the proposal includes a safety assessment than that will be included in the document analysis.

2.4 Analysis

The results from the micro-experiments and rule adaptation are tested for their effectiveness on facilitating the necessary conversations to reduce the gap between prescription and practice, changing rules to better match daily operations and break through the organisational complicity evident in the target organisation. The effectiveness of the new approach is determined by answering the following six questions that (in sequential order) serve the analysis:

1. was the intervention conducted as intended and prescribed?
2. will the intervention become part of daily routines?
3. which rules/procedures need/needed to be modified?
4. has work-as-imagined been permanently changed?
5. has the gap between work-as-imagined and work-as-done been reduced?
6. is the micro-experiment sufficient justification for adjusting rules or is more needed?

The answer to question 5 determines if the goal to reduce the gap has been achieved. The answers to questions 2, 3, and 4 show if the second goal (change the rules to better match daily operations) has been achieved. Finally, the answers to questions 1 and 6 show if the proposed method is sufficient for changing work-as-imagined and thus breaks through the organisational complicity of deviating from prescription.

3. Findings

3.1 Selection of cases

Four cases have been selected to be subjected to this new approach:

- Reported issues where goal conflicts occurred at the operational unit (N=37)
- Of these cases selected for further analysis and description (N=34)
- Of these, concerning gaps between work-as-done and work-as-imagined (N=22)
- Of these, of an enduring nature (N=9)
- Of these, cases that included an explicit desire of the target organisation's management to reduce the gap between work-as-done and work-as-imagined (N=5)
- Of these, the availability of employees to conduct the micro-experiments (N=4)

All four cases differ greatly in terms of impact on daily operations (from local to airbase-wide) and work-as-imagined (from local procedures to international flight regulations) and are briefly introduced in Table 1. In Appendix 1 more detailed information on the scenario's and interventions is provided.

3.2 Summary of the four micro-experiments

Table 1 provides a process summary of the four micro-experiments. Detailed information about the cases is provided in Appendix 1.

Table 1: Process summary of four micro-experiments

Micro-experiment Half-day for developing innovations	**Micro-experiment Automating flight planning**
March 7 to April 1, 2022	April 18 to May 13, 2022
6 participants of which 1 problem owner. Role participants: working on innovations half a day a week for 4 weeks. No changes in participants during intervention.	4 participants of which 1 problem owner. Role participants: for 4 weeks testing of two planning tools for effectiveness and speed. No changes in participants during intervention.
3 interviews conducted, at the start and end of the intervention, after the decision not to conduct a second intervention. Each interview lasted 1 hour. First interview: problem owner, LB, RJdB, assistant Second interview: problem owner, LB, assistant Third interview: problem owner, LB	Startup meeting with problem owner, RJdB and assistant. Interview with representative of MoD software team, LB and assistant, lasting 1 hour. Evaluation meeting option 1, with 2 participants, RJdB and assistant, lasting 1 hour. Evaluation meeting option 2, with problem owner, RJdB and assistant, lasting 1 hour.
No observations were made.	One observation of both tools tested, lasting 1 hour, with problem owner, RJdB and assistant.
Data collected: notes based on interviews, PPT presentation made by problem owner.	Data collected: notes based on interview, meetings and observation, email correspondence regarding intervention.
Role authors: conducting the interviews, facilitating the startup meeting, ensuring preconditions for intervention, analysing the results.	Role authors: conducting the interview, attending meetings, facilitating startup meeting, ensuring preconditions for intervention, analysing the results.

Micro-experiment Fire fighter status: hot refuel	**Micro-experiment Fire fighter status: picking up passengers**
May 31 to June 28, 2022	
13 participants of which 1 problem owner. Role participants: during observations and evaluation meeting discussing safety / risks of changed rules, implementing additional measures. No changes in participants during intervention.	8 participants of which 1 problem owner. Role participants: conducting safety assessment as a preparation to the intended intervention. No changes in participants during process.
Interview with problem owner (LB, RJdB, assistant), lasting 1 hour. Startup meeting with all participants, LB en RJdB, lasting 1 hour. Meeting with airbase commander, problem owner, LB, lasting 1 hour. Meetings after every demonstration with 10 participant including problem owner, LB and RJdB, lasting 30 minutes. Evaluation meeting with 13 participants, LB and assistant, lasting 1 hour.	Interview with problem owner (LB, RJdB, assistant), lasting 1 hour. Meeting with all participants to conduct safety assessment, lasting 3 hours. Meeting with 6 participants, including problem owner and representative of the MAA, lasting 1 hour. In depth interview with 3 representatives of the MAA.
6 observations, of which 3 demonstrations and 3 drills. The goal of the observations was to witness the impact of the changes on the risk perception of the participants.	No observations were made.
Data collected: notes based on attended meetings, interviews and observations and safety assessment.	Data collected: notes based on meetings, interviews, documents containing the correspondence between airbase and MAA, safety assessment.
Role authors: attending meetings, designing the intervention, acquiring permission, ensuring preconditions for intervention, analysing the results.	Role authors: attending meetings, analysing the collected data.

3.3 Execution of micro-experiments and rule adaptation

In this section we discuss the micro-experiments, rule adaptation and the analysis consecutively for each case.

3.3.1 Micro-experiment 1: Half-day for developing innovations

a) Gap between work-as-imagined and work-as-done

Operators are on a busy schedule of flight training. Every working hour needs to be accounted for (work-as-imagined). Since calendars are fully booked, there is no time to innovate. This needs to be scheduled officially because spending time on innovation without official acknowledgement feels like playing hooky (work-as-done).

b) Proposed intervention and micro-experiment

The intervention is creating dedicated innovation time for the military unit. Different interventions were considered, varying in length and frequency. Squadron management together with designated problem owner converged to the following proposal: for the duration of four weeks six operators would spend half a day a week on innovation. This means that six micro-experiments were conducted, each with the same design. The time reserved for innovation was protected from all other demands by squadron leadership.

c) Results of the intervention

The intervention was well received and received the full support of flight planning. The unique non-negotiable element made a big difference. Participants reported a greater sense of being part of the squadron's future, being able to provide input into work processes. Both participants and commanding officers reported a more structured, thoughtful and reasoned way of approaching innovations.

d) Proposed rule adaptation

As a result of the micro-experiment, it was concluded that only management desired allocating formal time to legitimise innovation. However, no new rules were needed, as it was concluded that innovation is a legitimate part of the squadron culture.

e) Analysis and result

The intervention was conducted as intended and prescribed. Operators will continue to apply the intervention and reserve time for innovation. However, it was felt that no rules or procedures need to be modified, because the micro-experiment showed that operators are able to dedicate enough time for innovation in non-allocated time without this having to be reserved in their schedules. This had been underestimated by management. Therefore, the gap between the work-as-imagined (that rules are needed to ensure innovation) and the work-as-done has been reduced: innovation is possible within the existing planning rules. The micro-experiments, which tested the necessity of forming new rules to ensure time for innovation, were instrumental in demonstrating this. The micro-experiment resulted in acknowledging that time can be spend on innovation without formalising this and without seeing this as playing hooky, which means that the approach was sufficient to break through organisational complicity.

3.3.2 Micro-experiment 2: Automating flight planning

a) Gap between work-as-imagined and work-as-done

For security reasons, every employee of the Defence organisation is supposed to use the software provided by the organisation, and not create their own tools (work-as-imagined). However, the formal tools are inadequate for efficient flight planning of the pilots. Therefore, a separate unofficial planning support tool has been developed by the pilots themselves and is used for flight planning.

b) Proposed intervention and micro-experiment

The intervention focused on speeding up the planning process and gaining an overview of who is available when. To this end, two different planning tools are tested: the official one provided by the defence organisation and the unofficial planning support tool. Both tools were tested by two operators for the duration of four weeks. This means two micro-experiments, one for each tool.

c) Results of the micro-experiment

As a result of the micro-experiment, it was concluded that the unofficial planning support tool was far superior to the one provided by the defence organisation in planning speed and oversight, saving 1-2 hours per week.

d) Proposed rule adaptation

All functionality of the unofficial tool should be introduced into the official tool, or the unofficial tool should be recognised and supported by the central organisation.

e) Analysis and result

The intervention was conducted as intended and prescribed. The micro-experiment has gone some way in convincing the defence organisation to improve the official tool. However, not all functionality of the unofficial tool has been incorporated. The planners still clandestinely use the unofficial tool for some aspects of their work, hiding this fact to ensure it is not taken away from them. Therefore, the gap between work-as-imagined (using only the provided software) and work-as-done has been partly reduced but not eliminated. The micro-experiments were sufficient justification for adjusting rules but budget, resource and security constraints have inhibited a complete alignment of work-as-done and work-as-imagined. Fearing the unofficial tool will be taken away, the continued clandestine use of it shows the chosen approach only partly broke through the organisational complicity on using own software.

3.3.3 Micro-experiment 3: Fire fighter status: hot refuel

a) Gap between work-as-imagined and work-as-done

Hot refuel means that aircraft are refuelled with the engines running. Because of an ignition risk, two fire trucks oversee a maximum of four aircraft refuelling at the same time. When two fire trucks are not available e.g. due to staff shortage, the number of aircraft to be refuelled is limited to two at a time. The additional safety risk of four (rather than two) aircraft refuelling with only one fuel truck in attendance seems very limited, and the constraint has a significant impact on the operational training calendar.

b) Proposed intervention and micro-experiment

The intervention focussed on changing this rule and to create a situation where four aircraft can be refuelled under the oversight of only one fire truck. Other solutions (varying number of supporting staff increases etc.) were also regarded but dismissed. In the design of the micro-experiment much time was spent on determining the maximum time needed for the fuel truck to reach the outmost aircraft. The micro-experiment consisted of three demonstrations and three fire alarm drills, conducted over a course of four weeks.

c) Results of the micro-experiment

The case for four aircraft being refuelled and the outmost catching fire were simulated several times. Parties identified additional measures to improve the safety case (e.g. restrict pedestrian access). The process of micro-experimentation initiated and facilitated a process of mutual dialogue and allowed stakeholders to develop the new procedure together.

d) Proposed rule adaptation

The proposal to allow four aircraft to be refuelled simultaneously with only one fire truck and accompanying evacuation procedures was accepted by the airbase commander on the basis of the (adapted) safety case. However, in a later stadium central command and airbase legal departments disputed the jurisdiction of the airbase commander to change these rules and at time of writing the formalisation of the new way of working is yet to be realised.

e) Analysis and result

The intervention was conducted as intended and prescribed. Our approach reduced the gap between work-as-imagined and work-as-done. Practice and the (changed) rules were aligned. Combining the perspectives, knowledge and expertise of all involved in the micro-experiment led stakeholders to discover additional risks not found in the previous risk assessment. The intervention resulted in all stakeholders agreeing to the rule changes and feeling safer than before. Work-as-imagined has permanently changed for all participants, although the legal department espoused objections after the intervention, and halted the formalisation of the new work practice. This means that at time of writing the new way of working is implemented, is described on paper, but not formalised, resulting in organisational complicity, which the research approach was unable to break through.

3.3.4 Micro-experiment 4: Fire fighter status: picking up passengers

a) Gap between work-as-imagined and work-as-done

The operating unit under study includes aircraft with the capacity to transport passengers. The number of fire trucks that are on standby when such an aircraft departs is dependent upon the number of people on board. When the number of fire trucks is limited (e.g. through staff shortage), passenger operations from the airfield are severely impacted. As a result, the aircraft depart without passengers from the airfield with fire trucks (in compliance with the rules), and then land in a nearby field to pick up their passengers where the airbase rules do not apply and without any fire trucks being available. This too is

compliant with the rules but seems to lessen safety barriers. As such, there is not a gap between practice and actual rules, although there is a clear gap with the intent of the rules.

b) Proposed intervention and micro-experiment

The intervention focusses on picking up passengers from the airbase with mitigations in place to compensate for the reduced fire fighter status with the intention an gaining a formal regulatory exception for this situation. However, the micro-experiment using the intervention did not commence because the military aviation authority (MAA) will not allow it.

c) Results of the micro-experiment

Not applicable.

d) Proposed rule adaptation

The MAA will not grant an exception because international regulations specify the fire fighter status for aircraft of this size, regardless of the number of passengers. The current practice was already an exception to these international rules. Having been made aware of the situation, the MAA has made several regulations regarding the availability of water to extinguish fires more stringent. Whereas the airbase considers the issue a problem with rules, rule management and exception management, the MAA considers it a problem with business operations (e.g. to recruit more supporting staff).

e) Analysis and result

Passenger pick-up from the nearby airfield remains a working practice. The gap between prescription and practice has not been reduced; in fact the gap has widened as a result of more stringent rules. On closer examination of the procedures for rule management and risk assessment, it was found that the military organisation can actually be noncompliant to the rules defined by the MAA as long as the appropriate authority takes responsibility for the risks involved, in this case the airbase commander. At time of writing, this course of action is still under discussion.

In depth interviews show the MAA distrusts the military organisation, accusing them of an inappropriate "can-do" mentality and looking for an easy way out instead of seriously solving problems with business operations. The MAA fears a slippery slope regarding exceptions. The fear and distrust that the MAA expressed prevents the adaptation of rules and impedes the reduction of the gap between work-as-imagined and work-as-done. In this case our approach was unsuccessful in facilitating the necessary conversations to reduce the gap between prescription and practice, changing rules to better match daily operations and break through the organisational complicity.

4. Discussion

In this study we aimed to identify whether a combination of a framework proposed by Hale and Borys (2013b) with micro-experiments (Dekker, 2018) facilitates the necessary conversations to reduce enduring gaps between work-as-imagined and work-as-done, enables changes to the rules to better match daily operations and to understand whether this method allows us to break through the

organisational complicity. We studied this through the observation of four cases in which it was desirable to lessen a gap between work-as-done and work-as-imagined. The framework was applied by identifying potentially successful interventions and testing these with micro-experiments. We monitored the proposals to initiate rule changes. The results from the micro-experiments and rule adaptation were tested for their effectiveness on facilitating the necessary conversations to reduce the gap between prescription and practice, changing rules to better match daily operations and break through the organisational complicity evident in the target organisation.

Our findings show that of the four cases, in only one case (half-day for developing innovations) the chosen approach was fully successful in facilitating the necessary conversations to reduce the gap between prescription and practice, changing rules to better match daily operations and break through organisational complicity. In two cases (automating flight planning and hot refuel) the approach was partially successful. In the first case there was a reduction in the gap between rules and reality, but still some organisational complicity remained evident. In the second case there was a reduction in the gap between rules and reality on paper, but other authorities dismissed this and so the new way of working has not been implemented. In the fourth case (picking up passengers) the approach was not successful in facilitating the necessary conversations to reduce the gap between prescription and practice, let alone changing rules to better match daily operations and break through organisational complicity. In the three cases we found that the process of micro-experiments were technically successful in showing the effectiveness of the interventions and generating shop floor commitment for them, as predicted by Dekker (2018) and De Boer (2021).

This research allows us to reflect on the definition of organisational complicity, originally described as supervisors ignoring violations of rules that hamper operations followed by a 'conspiratorial silence about the poor rules' (Hale & Borys, 2013a). The case of the picking up passengers shows that working according to rule, knowing that non-compliance is actually safer, and not being overtly transparent about it also obscures the need to modify rules. We suggest that these cases be included in the definition of organisational complicity.

Reducing gaps between work-as-done and work-as-imagined by modifying rules is a separate step that does not follow automatically from combining a procedural approach with micro-experiments, and was not explicitly described in Hale and Borys (2013b) and De Boer (2021). In the first, successful case no rule changes were needed and so was trivial in this regard. Reducing the gap by modifying the rules was impeded when the mandate for the rules lay outside the group involved in the micro-experiments and/or was distanced from the operational team, as evident in the cases of the planning tool, hot refuel and picking up passengers. These cases show that the desire to reduce a gap between rules and reality is in fact not strong and homogeneous enough throughout the organisation to overcome differences in perception between entities.

The common denominator in the three latter cases is fear. In the case of the planning tool, the planners fear the central department enough not to be transparent about their solution. In the case of the hot refuel, the legal department fears that the airfield commander does not have the mandate to change the ways of working, and so a stalemate occurs. In the third case, the MAA is feared by the operators so that they comply to the rules, knowing that non-compliance is actually safer. This fear and lack of trust, previously termed as the bogeyman belief system regarding non-compliance (Boskeljon-Horst, De Boer, et al., 2023), impeded the reduction of the gap between work-as-imagined and work-as-done. The fact

that a belief system gets in the way of actual change seems to have been overlooked in the framework proposed by Hale and Borys (2013b) and is insufficiently expressed in the micro-experiments proposed by Dekker (2018a). Furthermore, Hale and Borys (2013b) seem to have underestimated the effort it takes to bring rule makers and users together. Even if all the steps are followed, this still does not mean that one can achieve structural change in ill-functioning rules.

What seems to make a difference when trying to reduce the gap between prescription and practice and breaking through organisational complicity is the issue of conflicting values and whether or not these are resolved. All four cases featured conflicting safety values. While analysing the data of the micro-experiments, we serendipitously discovered a fifth case that is exemplary of conflicting values and a gap between work-as-done and work-as-imagined. This case confirms that the real effort may lie not so much in technical alignment but in enhancing restorative just culture, resolving value conflicts and, subsequently, preventing moral disengagement.

4.1 Crew endurance case

The serendipitous case is about crew endurance on tactical training missions. On training missions the entire unit (operators, maintenance, ground support, etc.) sleeps in tents, mimicking a situation close to the front of a conflict zone. With so many people in one tent, all with different working and sleeping hours, the crew lacks rest. Operators grow increasingly fatigued in the course of the mission and feel that this will impact their safety when flying mission. The operators began addressing this problem about five years ago, but they are not taken seriously – and nothing has changed since. The operators feel that the issue is being trivialised. To solve this matter, the operators have asked for rules on sleep in the operations manual dictating the sleeping arrangements and both quantity and quality of sleep. Recently however, they have been threatened with personal consequences for their careers if they do not drop the matter. This has resulted in a perception of psychological unsafety (Edmondson, 2019), which the operators call a 'culture of fear'.

4.2 Analysis of the crew endurance case

In discussing the case with the different parties, both individually and in group meetings (including airbase commander) it became clear that the real issue was not the sleeping arrangements but the different ideas and values on preparing for missions abroad and being a good pilot, possibly even a good soldier. The tactical training missions are conducted in the woods during the bad weather conditions of autumn or winter, in camps without any luxury, just the basics of food and sleeping arrangements. Management finds these conditions necessary to train command and control under duress and difficult circumstances, specifically, being able to execute flight missions while fatigued and in miserable training conditions. This contributes to management's wish for proper preparation for combat and is part of the value of a good pilot and a good soldier. For the operators, on the other hand, proper preparation for flight missions in combat means being able to execute every manoeuvre to perfection and thus creating a safety margin to fall back on. This means that if the circumstances become difficult, they can still perform well enough because their extensive training as automated part of their

skills and require less cognitive effort. For operators, the value of a good pilot and a good soldier means being an expert, which requires every flight training hour possible.

Neither management nor operators recognised the differences in these conflicting values. A discussion is impeded by a 'culture of fear'. Operators felt that from a certain rank up, commanding officers were reluctant to allow crew to speak up about difficult matters – usually because their own career is on the line. Metaphorically this behaviour is called the watermelon-syndrome: what is red on the inside – the operator level – becomes green on the outside – the level of airbase commander. Each subsequent level above operators slightly brightens the message, until a 'no' is turned into a 'yes' and commanding officers can show the hierarchy that everything is fine. For instance, during one tactical training exercise, pilot fatigue caused a near-miss incident, which the operators reported in the occurrence reporting system. However, commanding officers passed it on in the chain of command, without mentioning the operators' cause analysis of fatigue. During one tactical training exercise the operators *were* permitted to sleep in buildings instead of tents. According to the operators, this was possible because the exercise leader had no intention of pursuing a career in the military organisation and did not care how this decision would affect his image within his chain of command.

The crew endurance case makes explicit that a discussion on (safety) values is necessary to align rules with operational needs, but that this is difficult because the values are solidified in a belief system. Even starting the discussion is impeded by an implicit (or sometimes explicit) threat to career development (the bogeyman).

4.3 Comparision with the original cases

Although the crew endurance case was serendipitously discovered and not originally part of the designed approach of this research, we were still able to compare this case to the four micro-experiments due to the in depth interviews we conducted. Table 2 provides an overview of the comparison of the cases.

Table 2: Overview of the results of four micro-experiments and crew endurance case

Data source	**Research goals**	**Micro-experiment Half-day for developing innovations**	**Micro-experiment Automating flight planning**	**Micro-experiment Fire fighter status: hot refuel**	**Micro-experiment Fire fighter status: picking up passengers**	**Crew endurance case**
	Reducing gap	✓	✓	✓	X	Stale-mate
	Changing rules	✓	~	~	X	
	Break through o.c.	✓	~	~	X	
	Factors					
Desk research, interviews	Own mandate	✓	X	✓	X	✓
Observations, template	Trying ME	✓	✓	✓	X	X
In depth interviews	Distrust	Removed	Partly removed	Removed Added by legal d.	MAA distrusts mil.org.	Operators distrust management
Document review	Organizational compl.	X	~	Yes	Yes	Yes

The overview of the results show that having the mandate for rules helps, but is not sufficient for a successful micro-experiment and subsequently for reducing the gap between work-as-imagined and work-as-done, changing the rules and breaking through organisational complicity. We found that a situation in which distrust is removed, fear of noncompliance is not a factor (no bogeyman) and having mandate for rules, result in successful micro-experiments and therefore achievement of reducing gaps between work-a-imagined and work-as-done, changing rules and resolving organisational complicity. Having the mandate for rules but being unable to remove distrust among stakeholders and/or the fear of noncompliance (bogeyman) results in not being able to achieve as structural change (hot refuel and endurance case). This situation worsens if the mandate resides elsewhere (picking up passengers).

The crew endurance case can be compared to the hot refuel micro-experiment. In both instances the military organisation has the mandate for the rules involved. However, in the crew endurance case it was hard to get a conversation even started, whereas in the hot refuel micro-experiment the rules were adapted and all parties involved felt safer than before, even though the legal departments put a hold on the process. The difference between the two cases seems to be the ability to resolve conflicting safety values and a speak-up climate. All parties in the hot refuel micro-experiments were able to speak freely about the risks during the entire process and were therefore able to discuss the value of safety. The operators in the crew endurance case first felt ignored and later on became reluctant to discuss the issue because of insults and personal threats. The value of safety was never discussed, just the operators' – in the eyes of the chain of command – whingeing about sleeping in hotels.

Underpinning these results, in previous research it was found that the hierarchical structure of the military organisation hampered psychological safety and speaking up (Boskeljon-Horst et al., 2022) (Boskeljon-Horst, Snoek, et al., 2023). In those instances trust is lacking. Not speaking up can impede reducing the gap between work-as-done and work-as-imagined, as both automating flight planning and the crew endurance case show and enhance organisational complicity as the picking up passengers scenario shows. In contrast, in the hot refuel micro-experiment the speaking up done by the support departments that felt their safety was compromised, enhanced the reduction of the gap between practice and prescription. In this case, trust was restored.

Furthermore, in previous research (Boskeljon-Horst et al., 2022) it was found that a bogeyman belief intensifies the inhibition against speaking up. The shared experiences in the crew endurance case confirm the bogeyman belief: speaking up might end in personal consequences. The bogeyman also seems visible in the picking up passengers scenario. In this case, the MAA distrusts the target organisation and fears that a regulatory exception would create less compliance. In fact, the MAA's countermove, adding an extra restriction, tightens the existing rules. Although possible, noncompliance as a response to this countermove, is not an option, even though it would enhance safety.

Interestingly, in previous research it was found that operators think there are too many rules and procedures (Boskeljon-Horst et al., 2022) (Boskeljon-Horst, De Boer, et al., 2023). However, in the crew endurance case, because operators lacked trust in management, they wanted to add to the rules to feel protected. Usually, as observed in the picking up passengers scenario, rules are set because management does not trust operators (Antonsen et al., 2008). In the crew endurance case, operators preferred the classical paradigm on rule management (Hale & Borys, 2013a), considering the fatigue rule as carved in stone, adding to the promise of compliance found previously (Boskeljon-Horst, De Boer, et al., 2023). It was also found that management usually considers rules and compliance more

important than operators (Boskeljon-Horst et al., 2022). We saw this again in the innovation scenario, where squadron staff desired rules to govern innovation. However, in the crew endurance case management said that rules do not provide protection, reasoning that operators would not stick to them or change them if compliance was not possible, indicative of a constructivist paradigm (Hale & Borys, 2013a). This approach, however, added to the distrust operators already felt.

Overall, these results show that our approach, combining the research framework (Hale & Borys, 2013b) with micro-experiments (Dekker, 2018) is valuable in reducing the gap between paper and practice. That said, technical alignment is not in itself capable of overcoming the various obstacles in the way of actually changing rules and breaking through organisational complicity. Factors such as distrust and fear of noncompliance (bogeyman) not only hamper reducing gaps between paper and practice and changing rules, but actively enhance organisational complicity.

To be successful, it seems necessary to supplement this approach with a focus on the softer side of an organisation, that of trust and fear as seen in psychological betrayal and moral disengagement and countered by psychological safety, restorative just culture and psychological contracts.

It furthermore needs leaders throughout the entire chain of command to display moral courage to ensure that gaps between work-as-imagined and work-as-done are reduced and to actually change the rules if needed and break through organisational complicity, even if that means breaking the promise of compliance in cases where noncompliance is actually safer.

The highest management level of the Defence organisation needs to show the same behaviour it asks of its operators: to speak up about risks and obstructions, only in their case towards the press, public and politicians. This is needed to break the vicious circle of the classical paradigm and organisational complicity (Hale & Borys, 2013a) in which the focus on compliance, reinforced by a promise to press, politicians and public to remain compliant, creates a bogeyman belief system and tension throughout the organisation and reduces trust. The tension is due to conflicting goals, a retributive just culture and fear of consequences because of the hierarchical structure of the organisation lacks psychological safety (Boskeljon-Horst, Snoek, et al., 2023). The moral courage of leaders will enhance the effectiveness of tools like the micro-experiment and a research framework such as proposed by Hale and Borys (2013b).

4.4 Psychological contract, psychological and organisational betrayal and moral disengagement

To obtain insight into the processes of the crew endurance case as a means to explain the findings of the four micro-experiments, we briefly explore the topics of psychological contract, psychological betrayal, institutional betrayal and moral disengagement in relation to this particular case as well as the rest of this research. It is not the intention to provide a full overview of these topics. Just to cover the scientific underpinnings would require a separate paper. The topics are used only to see if a new light can be shed on the difficulty of changing rules

Besides a financial contract between employees and their organisation, which stipulates rights and duties, employees usually experience a psychological contract. A psychological contract implies the beliefs of an individual in a reciprocal relationship and commitment between the employee and the organisation. The employee has unwritten expectations of how the organisation will respond to their

behaviour (Rousseau, 1989). Violating a psychological contract means that the organisation did not respond as expected to the employee's contribution (Frey, 2018). This damages the relationship between the organisation and the employee (De Clercq et al., 2021). The trust between the two parties is broken in a form of trauma (Rousseau, 1989). Violation of a psychological contract is also called psychological betrayal (Chan, 2009). If the organisation fails to respond to recurring psychological betrayal this constitutes a larger systemic problem called organisational or institutional betrayal (Frey, 2018). If the organisation considers self-perseverance and self-interest more important than employee welfare, resulting in betrayal of the latter if the employee stands in the way of attaining these goals, this is called organisational egoistic betrayal (Chan, 2009). Psychological contracts and psychological or organisational betrayal are about values. When an employee does not keep their own or the organisation's norms and underlying values this is called moral disengagement (Bandura, 1999).

In a hierarchical organisation there is always some tension regarding speaking up within the chain of command (van Baarle & Blom-Terhell, 2022) and a hierarchal culture that makes speaking up more difficult (Sharygina-Rusthoven, 2019), increases the risk of moral disengagement (The Hague: Onderzoeksraad voor de Veiligheid, 2017). This can be overcome by psychological safety enhancing speaking up (Edmondson, 2019) and a restorative just culture focussed on repairing trust (Dekker, 2016). If moral disengagement is not dealt with it this can, in extreme situations, result in moral injury as a consequence of violating the values and the beliefs about right and wrong (van Baarle & Molendijk, 2021; Brewer, 2021; Litz et al., 2009).

Applying this theory to the crew endurance case seems to show the following. Operators are continuously confronted by the idea of speaking up throughout their career. The belief is that risks spoken up about will be dealt with seriously (psychological contract). Speaking up is a core value of the military organisation. In the case of crew endurance, the organisation, however, did not respond as expected. In fact, the risk was downplayed, the operators were denigrated with insults and silenced by using personal consequences as threats. This broke the psychological contract on speaking up, a situation compounded by the conflicting norms and values regarding 'good pilot/soldier'. The operators felt betrayed by the organisation. The organisational betrayal is visible in the fact that this risk has been a recurring theme in the past five years and on every single account the operators felt ignored. One could even see signs of organisational egoistic betrayal because it is in the airbase's interest to put operators in tents as they are cheaper than hotels. Allocating budget to hotel rooms means taking budget away from other – in the view of airbase leadership – more pressing matters. Furthermore, *not* allocating budget is consistent with management norms and values for 'good pilot/soldier'.

After five years of trying, the operators gave up on the matter, no longer spoke up about the risks they encountered because of the belittling comments and their fear. Signs of moral disengagement began to appear. For the operators under threat, the effect of broken trust was so profound that they have said they no longer trust the organisation on any matter. Previous research shows that the military organisation lacks a restorative just culture; it is predominantly retributive (Boskeljon-Horst, Snoek, et al., 2023). According to the operators it also lacks psychological safety. Moral injury might occur if this case was not dealt with seriously.

4.5 Contribution

This study contributes to the safety literature by confirming the value of Hale and Borys' framework (2013b) with micro-experiments (Dekker, 2018) for proving the effectiveness of interventions and generating the necessary shop floor commitment to realise them. It further contributes by demonstrating the hiatus in current literature on the steps following the choice of an intervention to reduce the gap between work-as-done and work-as-imagined by modifying rules. We have identified the impeding factors.

We contribute to the existing ideas for changing rules (Hale and Borys' framework, micro-experiments) by adding the factors of (conflicting) values, trust and bogeyman. We contribute to the safety literature by expanding the definition of organisational compliance as described in Hale and Borys (2013a). This expansion seen in the fear of noncompliance, reminiscent of a bogeyman, resulting rule following whereass local ingenuity or noncompliance would be safer, points more clearly to the risks of the classical approach to rule management, which is underreported by Hale and Borys (2013a).

4.6 Future research

To achieve real progress in enhancing both physical and psychological safety in an organisation, we recommend that future research focuses on such topics as psychological safety and organisational betrayal and combine these with such topics as compliance and adaptation to find supplementary explanations for the enduring differences between work-as-imagined and work-as-done.

5. Limitations

This study has definite strengths. The availability of operators and stakeholders to conduct micro-experiments as well as the support in the chain of command greatly strengthened the micro-experimental process. Our qualitative methodology of observations combined with interviews provided valuable insights into the experiences of participants and helped articulate the patterns found. The use of both inside and outside researchers not only aided data collection, since the inside researcher knows the population, the customs and language used, but also enhanced objectivity as the outside researcher could question the findings as an antidote to any possible blind spots in the observations of the insideresearcher.

However, the scope of this study was limited to one particular command in the RNLAF and, within this command, to the enduring gaps between work-as-imagined and work-as-done that the operators encountered. Although the sample of operators resembles other operator squadrons in RNLAF commands, it is less representative of the non-operator part of the RNLAF. That makes the study results directly applicable to other operator squadrons in the RNLAF but less so elsewhere, both within and outside the RNLAF, or even outside the Defence organisation organisation.

Added to this is the low number of micro-experiments. This study focused on micro-experiments in four cases, of which the MAA stopped one before it could start. Nonetheless, the reason for not doing the

micro-experiment presented data because it showed how difficult it actually is to even suggest changing rules for which one has no mandate. The cases differed greatly in content and existing rule base. However, our combined methodology produced comparable results. This seems to indicate some form of data saturation for this albeit limited number of micro-experiments.

6. Conclusion

We found that the micro-experimental process was technically successful at showing the effectiveness of the interventions, generating shop floor commitment for them, and made it possible to reduce the gaps between work-as-done and work-as-imagined. However, we also found that actually modifying rules was impeded by the tension between compliance and adaptation and the presence of a bogeyman belief regarding noncompliance. Our study confirms Dekker (2018) and De Boer's (2021) suggestion that micro-experiments are effective interventions and a good means to gain shop floor commitment, but that the difficulty of actually modifying rules and breaking through organisational complicity has been overlooked.

This study also shows that the chosen framework and micro-experiments only partly end the bogeyman belief. The bogeyman reappears the moment rules need to be modified. A mere technical solution, as proposed by Hale and Borys (2013b) and De Boer (2021), will not suffice if the tension between compliance and adaptation remains unsolved and the bogeyman is kept alive. The approach used in this study was insufficient to break through organisational complicity. Adding focus on such subjects as psychological safety, restorative justice, psychological contracts and betrayal, and moral disengagement seems needed.

This means that enhancing safety requires more than a technical solution to close the gap between work-as-imagined and work-as-done and may need even more than a paradigmatic change, from the classical paradigm to the constructivist paradigm. It seems to require expanding the constructivist paradigm with knowledge of these topics. Future research that combines these scientific domains is recommended to understand the underlying patterns and to achieve real progress in aligning paper and practice.

References

Antonsen, S., Almklov, P., & Fenstad, J. (2008). Reducing the gap between procedures and practice - lessons from a succesful safety intervention. *Safety Science Monitor, 12*(1). https://www.researchgate.net/publication/229005586

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*(3), 193–209. https://doi.org/10.4324/9781315799292-3

Besnard, D., & Greathead, D. (2003). A cognitive approach to safe violations. *Cognition, Technology & Work, 5*(4), 272–282. https://doi.org/10.1007/s10111-003-0131-1

Bieder, C., & Bourrier, M. (2013). *Trapping safety into rules: How desirable or avoidable is proceduralization*? Ashgate Publishing Co. https://doi.org/10.1080/00140139.2013.819675

Boskeljon-Horst, L., Boer, R. J. De, Sillem, S., & Dekker, S. W. A. (2022). Goal Conflicts , Classical Management and Constructivism : How Operators Get Things Done. *Safety, 8*(37).

Boskeljon-Horst, L., De Boer, R. J., Steinmetz, V., & Dekker, S. W. A. (2023). Aircrews, rules and the bogeyman: Mapping the benefits and fears of noncompliance. *Safety, 9*(1).

Boskeljon-Horst, L., Snoek, A., & Baarle, E. van. (2023). Learning fom the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force. *Safety Science, 161*.

Brewer, K. C. (2021). Institutional betrayal in nursing: A concept analysis. *Nursing Ethics, 28*(6), 1081–1089. https://doi.org/10.1177/0969733021992448

Chan, M. L. E. (2009). "Why Did You Hurt Me?" Victim's Interpersonal Betrayal Attribution and Trust Implications. *Review of General Psychology, 13*(3), 262–274. https://doi.org/10.1037/a0017138

Christensen, C. M., Hall, T., Dillon, K., & Duncan, D. S. (2016). Know Your Customers' "Jobs to Be Done." *Harvard Business Review, September*.

Damoiseaux-Volman, B. A., Medlock, S., van der Eijk, M. D., Romijn, J. A., Abu-Hanna, A., & van der Velde, N. (2021). Falls and delirium in older inpatients: Work-as-imagined, work-as-done and preferences for clinical decision support systems. *Safety Science, 142*. https://doi.org/10.1016/j.ssci.2021.105355

De Boer, R. J. (2021). *Safety Leadership. A different, doable and directed approach to operational improvements*. CRC Press.

De Clercq, D., Azeem, M. U., & Haq, I. U. (2021). But they promised! How psychological contracts influence the impact of felt violations on job-related anxiety and performance. *Personnel Review, 50*(2), 648–666. https://doi.org/10.1108/PR-07-2019-0388

Dekker, S. W. A. (2003). Failure to adapt or adaptations that fail: Contrasting models on

procedures and safety. *Applied Ergonomics, 34*(3), 233–238. https://doi.org/10.1016/S0003-6870(03)00031-0

Dekker, S. W. A. (2014). *Safety Differently. Human factors for a new era*. Apple Academic Press Inc.

Dekker, S. W. A. (2016). *Just culture: Restoring Trust and Accountability*. CRC Press Taylor & Francis Group.

Dekker, S. W. A. (2018). *The Safety Anarchist*. Routledge Taylor & Francis Group.

Dul, J., & Hak, T. (2007). *Case study methodology in business research*. Routledge

Edgett Collins, S. (2007). Criminalization of negligence in nursing: A new trend? *The Florida Nurse, 55*(1), 28.

Edmondson, A. C. (2019). *The fearless organization. Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth*. John Wiley & Sons, Inc.

Frey, L. L. (2018). When it Hurts to Work: Organizational Violations and Betrayals. *New Directions for Teaching and Learning, 153*, 87–98. https://doi.org/10.1002/tl.20284

Gauthereau, V., & Hollnagel, E. (2005). Planning, control, and adaptation: A case study. *European Management Journal, 23*(1), 118–131. https://doi.org/10.1016/j.emj.2004.12.016

Halbesleben, J. . B., Wakefield, D. ., & Wakefield, B. . (2008). Work-arounds in health care settings: Literature review and research agenda. *Health Care Management Review, 33*(1).

Hale, A., & Borys, D. (2013a). Working to rule, or working safely? Part 1: A state of the art review. *Safety Science, 55*, 207–221. https://doi.org/10.1016/j.ssci.2012.05.011

Hale, A., & Borys, D. (2013b). Working to rule or working safely? Part 2: The management of safety rules and procedures. *Safety Science, 55*, 222–231. https://doi.org/10.1016/j.ssci.2012.05.013

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. *Paper Presented at the Gestaltung Nachhaltiger Arbeitssysteme, 58. Kongress Der Gesellschaft Für Arbeitswissenschaft*, Pp.19-24.

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Litz, B. T., Stein, N., Delaney, E., Lebowitz, L., Nash, W. P., Silva, C., & Maguen, S. (2009). Moral injury and moral repair in war veterans: A preliminary model and intervention strategy. *Clinical Psychology Review, 29*(8), 695–706. https://doi.org/10.1016/j.cpr.2009.07.003

Loukopoulos, L. D., Dismukes, K., & Barshi, I. (2009). *The multitasking myth: Handling complexity in real-world operations*. Ashgate Publishing Limited.

Moppett, I. K., & Shorrock, S. T. (2018). Working out wrong-side blocks. *Anaesthesia, 73*(4),

407–420. https://doi.org/10.1111/anae.14165

Morrison, E. W. (2006). Doing the job well: An investigation of pro-social rule breaking. *Journal of Management, 32*(1), 5–28. https://doi.org/10.1177/0149206305277790

Plioutsias, A., Stamoulis, K., Papanikou, M., & de Boer, R. J. (2020). Safety differently: A case study in an Aviation Maintenance-Repair-Overhaul facility. *MATEC Web of Conferences, 314*, 01002. https://doi.org/10.1051/matecconf/202031401002

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate.

Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal, 2*(2), 121–139. https://doi.org/10.1007/BF01384942

Sharygina-Rusthoven, M. O. (2019). *The Role of Context in Proactive and Voice Behavior*. PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam.

Shorrock, S. (2016). *The Varieties of Human Work*. https://humanisticsystems.com/2016/12/05/the-varieties-of-human-work/

Snowden, D. J., & Boone, M. E. (2007). A Leader's Framework for Decision Making. *Harvard Business Review, November*.

The Hague: Onderzoeksraad voor de Veiligheid. (2017). Mortierongeval Mali.

Tucker, A. L. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management, 22*(2), 151–169. https://doi.org/10.1016/j.jom.2003.12.006

Tucker, A. L., Zheng, S., Gardner, J. W., & Bohn, R. E. (2020). When do workarounds help or hurt patient outcomes? The moderating role of operational failures. *Journal of Operations Management, 66*(1–2), 67–90. https://doi.org/10.1002/joom.1015

van Baarle, E., & Blom-Terhell, M. (2022). The Roof, the Roof, the Roof is on Fire, Moral Standards and Moral Disengagement in Military Organisations. In *Ethics and Military Practice* (pp. 24–39). Brill.

van Baarle, E., & Molendijk, T. (2021). Resilience as the Road to Mental Readiness? Reflections from an Ethics-of-care Perspective. *Journal of Military Ethics, 20*(2), 129–144. https://doi.org/10.1080/15027570.2021.1973721

Appendix 1

Intervention 1: Half-day development

Problem, impact, intervention and goal

The squadron wants to innovate but is unable to turn intention into action. Short-term work pressure usually wins over long-term innovative achievement. The intervention created time for innovation by reserving half a day a week in the training calendar for participants to develop creative ideas which potentially could solve frustrations and obstacles in the working environment.

Duration, participants, stakeholders, conditions

From March 7 to April 1, 2022, six operators not on study or holiday leave spent half a day a week on developing creative ideas for innovation. This half-day was not allowed to be withdrawn due to operational necessity. No other stakeholders were identified besides the target organisation. If any unforeseen situations arose, the authority to terminate the intervention would reside with the commanding officer of this squadron. Before the intervention began, the commanding officers informed the entire squadron about the intervention. Flight planning were not allowed to deploy the participants cleared for this intervention.

Gap between work-as-imagined and work-as-done

Operators are on a busy schedule of flight training. Every working hour needs to be accounted for (work-as-imagined). Since calendars are fully booked, there is no time to innovate. This needs to be scheduled officially because spending time on innovation without official acknowledgement feels like playing hooky (work-as-done).

Results of the intervention

The intervention was well received and received the full support of flight planning. The unique non-negotiable element made a big difference. Participants reported a greater sense of being part of the squadron's future, being able to provide input into work processes. Both participants and commanding officers reported a more structured, thoughtful and reasoned way of approaching challenges. When developing ideas, participants found it difficult to identify which departments or people should be approached for the problem at hand.

The impact on flight management was minimal, only three infringements to flight planning, which could be solved in two cases without breaching the half-days for the intervention. However, planners felt it was unrealistic for all operators to be scheduled half a day off per week as this was inconsistent with annual flight-hour targets.

Six operators were released from flight planning and given the space to think up new ideas. One pilot was given an existing assignment to work on because he was unable to come up with an idea of his own. By the end, six ideas had emerged, of which two have already been partially worked out. Squadron management has seen little result (output) from this intervention and is therefore slightly sceptical but recognises the value of ownership of the own working environment at operator level.

Because of these initial results, after the intervention ended, the squadron intended to initiate a second, slightly different intervention. Over a period of three months, eight operators would devote half a day a month to come up with imaginative ideas to improve the working environment and

develop them into a proposal to be discussed at the next flight commanders' meeting. Just before the start, the participants decided not to do the second intervention because they realised that time for innovation was not a problem for operators, only for squadron management. So, the project owner decided to look for ways to perpetuate the already innovative culture; a culture where in consultation they "sometimes safely ignore existing rules to keep innovating and improving." The biggest challenge is finding the right person with whom to share the innovative, creative solution

Six questions

Was the intervention conducted as intended and prescribed?
Yes
Will the intervention become part of daily routines?
Yes, the target oranisation will continue making time for innovations.
Which rules/procedures need/needed to be modified?
None, not needed
Has work-as-imagined been permanently changed?
Yes, management thought rules were needed. They now agree they are not needed.
Has the gap between work-as-imagined and work-as-done (theoretically) been reduced?
Yes, because no rules are needed, aligning work-as-imagined with work-as-done.
Is the micro-experiment sufficient justification for adjusting rules or is more needed?
Sufficient

Intervention 2: Automating flight planning

Problem, impact, intervention and goal

Due to a lack of pilots, it takes considerable work to plan flights. Any changes due to technical, meteorological or personnel reasons add to the workload. This intervention focused on speeding up the planning process and gaining an overview of who is available when.
The experiment tested two options. The first used an existing tool in the planning software that displays the operators' personal calendars beneath the flight plans. This option required flight members to track their personal schedules in the software's calendar.
Option two used a new scheduling tool, named after its inventor, Roel, an operator in the squadron. Using Excel, he created Roeltool v2.0 to first sort the unit's Outlook calendar items by name, then extract flight blocks from the planning software and display them similarly to option 1. This option required more work for planners to synchronise the tool with the planning software and Outlook, but the operators only needed to update their own calendars.

Duration, participants, stakeholders, conditions

From April 18 to May 13, 2022, two teams of two operators tested both options. To determine effectiveness, both were compared to the planning routine using the planning software without the calendar tool. Stakeholders were the squadron and Defence team responsible for planning software. At the beginning of the intervention, flights received information on either option 1 or 2. The flight members who tested option 1 were asked to fill out two calendars (in the planning software and their regular Outlook calendar) for the duration of the intervention. It was made sure that both tools worked as promised. If for some reason either one of the options did not work as expected, the flight had to report this to the problem owner first who would then decide whether or not to terminate the intervention. It saves time if planning changes can be processed clearly, quickly and effectively and if conflicts between calendar and practice remain a minimum the tool is considered a success.

Gap between work-as-imagined and work-as-done

For security reasons, every employee of the Defence organisation is supposed to use the software provided, not create their own (work-as-imagined). Thus, the Roeltool that operators were using before the intervention (work-as-done) deviated from work-as-imagined.

Results of the intervention

Option 1 worked reasonably well when Outlook capability was added to the secure planning software to enhance user-friendliness. Transferring personal calendars to the planning software was a prerequisite. It proved possible in the trial but was severely limited by the need to log into the secure defence network. Outlook, on the other hand, can be accessed from one's own phone or tablet. Provided the personal calendar was kept up to date, last-minute changes were a matter of seconds (used to be 10 minutes). Weekly planning, on the other hand, took nearly twice as long, 2-3 instead of 1-2 hours. This was initially due to the lack of multiple windows in the planning software, but a workaround was found for that. Time spend on weekly planning can be reduced if the interface tool in the planning software is improved. With the current version of the planning software the team prefers the former planning routine.

Option 2 (Roeltool) was fast. Updates to Roeltool can be made quickly as long as Roel is available to the squadron. Knowledge of the planning software or a planning software expert is necessary. The

advantage of Roeltool is that it can filter specific flights and it makes it easy to connect flights to people (or vice versa). It provides a clear overview of the planning and matches the normal operations of the squadron. Operators do not need to log in, but only enter their (changed) plans in Outlook. One still needs to use the Outlook calendar as intended. The software considers reminders put in Outlook as unavailability. Currently it is not possible to see two planned appointments simultaneously, but this does not interfere with the use of Roeltool. It is also useful for non-flying tasks as it is easy to see who is on leave. It takes 20 seconds to synchronise the calendars of 30 operators. A last-minute change takes a few seconds. Roeltool requires hardly any time for weekly planning and thus provides time savings of 1-2 hours per week. The team is eager to continue working with Roeltool. It is stable, saves a lot of time, and there is a user manual. The downside is that it makes the squadron dependent on Roel and there is no development manual, although there are comments in the code. And the intervention participants fear that the organisation will halt this development if they were to know about it. Nevertheless, the squadron decides to continue using Roeltool, hoping that one day it will be officially integrated in the planning software.

After the intervention, there was a meeting with the person in charge of planning software because continued MoD support is needed. This person already knew of Roeltool and adopted a part of it in the planning software. The new tool needs to be tested by the squadron. If everyone is satisfied with it, it will be distributed widely. Next step is connect Microsoft Outlook and the planning software.

Six questions

Was the intervention conducted as intended and prescribed?
: Yes

Will the intervention become part of daily routines?
: Partly; the software provided has been changed. For the changes not realised yet the operators will continue using the Roeltool.

Which rules/procedures need/needed to be modified?
: None, with the modification of a part of the Roeltool in a tool within the planning software the operators continue to use the software the RNLAF provides.

Has work-as-imagined been permanently changed?
: Yes, part of the Roeltool has been added to the official software provided.

Has the gap between work-as-imagined and work-as-done (theoretically) been reduced?
: Yes, because the provided software now matches practice better than it did before.

Is the micro-experiment sufficient justification for adjusting rules or is more needed?
: More is needed because the official software still does not fully match the practice of operators

Intervention 3: Fire fighter status: hot refuel

Problem, impact, intervention and goal

Hot refuel means that aircraft are refuelled with the engines running. Because of an ignition risk, either in the fuel depot or at the end of the hose inside the aircraft's fuel tank, two fire trucks oversee the maximum of four aircraft refuelling at the same time. Due to employee shortage in the fire department two fire trucks are not always available, reducing the maximum to two aircraft refuelling simultaneously. This has an impact on the operational training calendar of the target and other operational squadrons on the airbase. The airbase sets the rules for hot refuelling so it is in the authority of the airbase commander to change them. The intervention focussed on changing this rule to create a situation where four aircraft can be refuelled under the oversight of one fire truck.

A safety assessment had been conducted to identify the risks, mitigating measures and risk left after the mitigating measures. Although the airbase command has the authority to accept risks, multiple stakeholders strongly opposed the intention to change the rules, stating that reducing the number of fire trucks for an increasing number of refuelling aircraft would present more safety risks for the refuellers. The participation council had challenged the airbase commander on the matter.

Although the airbase commander could sign off the intended change, we asked him not to do so yet, but to agree to a demonstration and practice phase. We conducted a meeting with all stakeholders (operators, fire fighter department, air traffic control, refuelling department, safety department, project leader, Works Council) and listened to all their objections, needs and wishes.

Duration, participants, stakeholders, conditions

From May 31 to June 28, 2022, three demonstrations took place to test how long it would take a fire truck to reach the furthest helicopter. Then came three fire alarm drills involving a fire truck responding to a simulated fire, while employees safely left the fuel pads following a new evacuation plan. Each demonstration and drill tested the safety assessment in practice and all stakeholders evaluated each drill. They found no reasons to terminate the intervention prematurely. At the final evaluation meeting, all the stakeholders discussed and interpreted the findings. The airbase commander was informed about the outcome and related recommendations.

Gap between work-as-imagined and work-as-done

Flight hours need to be spent as efficiently as possible to ensure that all necessary training and assignments are accomplished (work-as-imagined). Airbase rules stipulate that it requires two fire fighter trucks to supervise the of use all four refuelling platforms simultaneously, This is a waste of flight time (work-as-done).

Results of the intervention

Almost everyone in the refuelling department has witnessed the fire brigade turning out and is impressed by how fast a fire truck can get to a helicopter to put out the fire. This has lessened the fear that safety would be compromised if only one fire truck supervised four aircraft. At the evaluation meetings, the stakeholders identified several issues that had not been considered when the safety assessment was conducted. Some issues were: marking the evacuation path, including a mowing plan because the path is partly on grass, which in the summer exposes employees to ticks; enhancing the visibility of signals in the dark to make sure the fire fighters notice an incident; distributing the new

procedure throughout the airbase and making sure demonstrations for new personnel will be continued in the future.

The three practice drills went a little differently than planned due to technical issues with the designated aircraft. No other aircraft were available, so the first drill had to be cancelled. However, the other two drills were successful. Participants discovered various learning issues, such as the right way to signal fire, visibility for both signalling and people evacuating in the dark, the evacuation direction and route, the placing of wheel blocks, and the decision to stay in the aircraft instead of evacuating. They also decided that monthly training drills would be necessary for the first year and after that, quarterly training. Again, all the stakeholders felt comfortable and safe with the new refuelling procedure.

At the final evaluation all stakeholders indicated they found the intervention process pleasant and appreciated the value of mutual dialogue and being able to finalise this changed procedure together. In fact, they decided to conduct these interdisciplinary meetings twice a year to discuss an issue or a problem and take advantage of all the different perspectives and expertise.

They also concluded that the changed procedures needed some adaptations after the intervention. For instance, before the intervention (but after the safety assessment) the procedures said that the operator of the aircraft always needed to leave the aircraft in case of a fire. But in some instances that might mean being blown over by a powerful jet of water. If the fire can be put out quickly, it might be safer to stay in the aircraft. The procedures were changed again, leaving this decision to the operator's discretion.

After the final evaluation, the mitigating measures were added to the safety assessment, the procedures were finalised and officially published for airbase personnel.

Six questions

Was the intervention conducted as intended and prescribed?
: Yes

Will the intervention become part of daily routines?
: Yes, the intervention has become a daily routine

Which rules/procedures need/needed to be modified?
: The airbase rules regarding the number of fire trucks needed to oversee four aircraft hot refuelling at the same time as well as the procedures regarding evacuating the refuelling pads.

Has work-as-imagined been permanently changed?
: Work-as-imagined has permanently changed for all participants, although the legal department espoused objections after the intervention.

Has the gap between work-as-imagined and work-as-done (theoretically) been reduced?
: Yes, practice and the (changed) rules were aligned

Is the micro-experiment sufficient justification for adjusting rules or is more needed?
: More is needed, the legal departments (temporarily) halted the formalisation of the new work practice.

Intervention 4: Fire fighter status: picking up passengers

Problem, impact, intervention and goal

Given the employee shortage in the fire department, fire fighter status fluctuates. The required fire fighter status for aircraft to be allowed to fly is based on size of the aircraft and number of persons on board (POB). The fire fighter status is denoted with letters and numbers. For instance, H3Cat7 means that aircraft belonging to category 7 are allowed to fly if the status is at level 3, meaning that three fire trucks are available in case of need. One level below is H2Cat4, indicating two fire trucks are available, allowing aircraft belonging to category 4 to fly. Fire fighter status rules are part of airbase rules, both civil and military. One required task is picking up troops for transport elsewhere. This is only allowed when fire fighter status is H3. However, due to the lack of employees resulting in H2 status, planned troops pick-up cannot be done some five times a year. To still get the job done, the aircraft pick up the troops at a neighbouring airfield, about 700 metres away, where airbase rules do not apply. However, there is no fire fighter support on this airfield.

The intervention focusses on picking up passengers from the airbase with mitigations in place to compensate for the reduced fire fighter status.

Duration, participants, stakeholders, conditions

Together with the RNLAF stakeholders (operators, fire fighters, air traffic control, safety officers), we used the RNLAF risk matrix to assess the current situation (pickups on a nearby airfield and airbase pickups with reduced fire fighter status) and devise mitigating measures. The matrix shows who has the authority to accept a risk and determines the impact of mitigating measures on the original risk. According to procedure, risk assessment is necessary when applying for a regulatory exception from the Military Aviation Authority (MAA). The intention is to test the regulatory exception for one year. Mitigations, like ceasing all other flying activities during landing and take-off of passenger aircraft, ensure the safety of the micro-experiment. Each passenger pick-up with reduced fire fighter status will be evaluated. The airbase commander will terminate the micro-experiment if the safety risk exceeds expectations.

Gap between work-as-imagined and work-as-done

Because flight operations face the inherent risk of crashing, when fire fighter assistance is vital for the survival of all persons on board, the safest way to conduct the most dangerous flight phases (taking off and landing) is with fire fighters present (work-as-imagined). Due to the shortage of fire fighters, passenger aircraft are not allowed to take off or land from the airbase so this is done on a nearby airfield without fire fighter support (work-as-done).

Results of the intervention

At time of writing the intervention still has to commence. The risk assessment has been sent to the MAA, along with the application for regulatory exception. The MAA replied that it will not grant an exception because international regulations specify the fire fighter status (H3) for aircraft of this size, regardless of the number of passengers. When the Netherlands ratified this regulation, it indicated that it would apply an exception for this particular aircraft (helicopter). The Dutch exception states that this aircraft will fly with reduced fire fighter status (H2) while maintaining a maximum of six persons on board.

This application for regulatory exception is actually an extension of an already granted exception which, the MAA has said, will continue in the future. However, the MAA has added the condition that the available quantity of extinguishing water must be of the same as prescribed for H3 protection, as this quantity is determined by the dimensions of the aircraft. According to the MAA, even with the new mitigating measures applied, an acceptable level of security cannot be maintained. Therefore, the general principle remains: operations involving more than seven persons on board must be secured in accordance with H3. Not only has the application for regulatory exception denied; the regulations in force have been made more stringent.

Conflicting perspectives

Interviews with the airbase commander and operators reveal that the airbase and the MAA have a different take on what the problem is. The airbase considers the issue a problem with rules, rule management and exception management. For them, the rules hamper task execution. The MAA considers it a problem with business operations. Not having enough fire fighters should not be solved by applying for a regulatory exception. It should be solved by hiring more fire fighters. Providing a regulatory exception is taking the easy way out, as it avoids having to deal with the actual problem. The MAA fears that if this regulatory exception is granted, the organisation will ask for another exception next time a business operation becomes a problem.

Six questions

Was the intervention conducted as intended and prescribed?
No, the intervention did not take place as it was rejected by the MAA.
Will the intervention become part of daily routines?
No, at this point, noncompliance is not an option.
Which rules/procedures need/needed to be modified?
The rules regarding fire fighter status when transporting passengers
Has work-as-imagined been permanently changed?
Yes, the existing rule base has been made more stringent.
Has the gap between work-as-imagined and work-as-done (theoretically) been reduced?
No, the gap has been widened as a result of a more stringent rule base.
Is the micro-experiment sufficient justification for adjusting rules or is more needed?
More is needed, as the MAA fears a slippery slope regarding exceptions

Chapter 7. Discussion and conclusions

Discussion and conclusions

1. Introduction

Understanding the relationship and gaps between work-as-done and work-as-imagined can enhance the safety and effectiveness of operational organisations. Being honest about this requires us to interrogate and possibly change belief systems throughout the organisation. What is really responsible for making things happen safely on the front line, despite conflicting goals, compliance demands and resource constraints? Discovering any of this, in turn, requires us to introduce concepts such as psychological safety and restorative just culture.

This thesis reports on five studies conducted in the Royal Netherlands Air Force (RNLAF), touching on safety culture, just culture, and compliance and adaptation. Together, the findings (and conceptual and theoretical implications) begin to provide an explanation of how an organisation achieves operational safety in a way that differs from previous beliefs that have tried to account for that.

The studies also illuminate the nature of, and reasons for, organisational-political stalemates on the gaps between work-as-imagined and work-as-done. Such stalemates assure silence throughout the organisational hierarchy on the 'open secrets' of how work actually gets done, preserving the interests of all stakeholders, as long as things do not go wrong. If and when they do go wrong, the RNLAF, like other organisations, has explanatory scripts at the ready that it can easily unfold by blaming a mishap on the errors, violations or non-compliance of frontline operators. The organisation thus completes another cycle of non-learning and secrecy about risks and retains the status quo of persistent gaps between work-as-imagined and work-as-done.

In this final chapter I answer the main research question, dealing first with the five sub-questions. Then I go on to reflect on the findings, relating them to the central research question, and finally I present my conclusions and recommendations for future research.

2. Research questions

This dissertation addresses the main research question:

How can the safety and effectiveness of the RNLAF be described and enhanced?

Subsequently it examines five sub-questions:

1. What is the predictive value of a safety culture assessment regarding the accident-proneness of an operational squadron? (Chapter 2)
2. What are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013a)? (Chapter 3)

3. Why do differences between the way work is done and prescribed endure despite their undesirability? (Chapter 4)
4. What complexities arise when attempting to foster a restorative just culture in a mostly retributive organisation? (Chapter 5)
5. Are micro-experiments enough to resolve the enduring gaps between work-as-imagined and work-as-done? (Chapter 6)

3. Main findings

Sub-question 1: What is the predictive value of a safety culture assessment regarding the accident-proneness of an operational squadron? (Chapter 2)

To answer this question we compared the results of a safety culture assessment (SCA) with the findings of an incident investigation. Described in Chapter 2, this comparison showed that the SCA and incident investigation differed in their results and descriptions of the maturity of the safety culture. This led to the conclusion that the SCA has limited predictive value for safety culture during an accident. In fact, the SCA inversely predicted the outcome. The investigation committee found that factors that the SCA considered indicative of a mature, proactive and generative safety culture were actually determined as the cause of this particular accident.

This study replicated the findings by Antonsen (2009) who compared a quantitative SCA with a 5-point Likert scale to an accident report. In contrast, we used an SCA with detailed workplace descriptions, enhancing the nuance and therefore perhaps the predictive value. However, despite the enhancements, our study reinforced the conclusion Antonsen drew on the limited predictive value of safety culture maturity assessments. If an organisation wants to predict safety outcomes, it will need a different measure than a maturity SCA needed.

Sub-question 2: What are the differences with regard to goal realisation when applying two conflicting paradigms regarding rule perception and management as described by Hale & Borys (2013a)? (Chapter 3)

To study these differences, we collected more than 30 scenarios from an operational unit of the RNLAF, as described in Chapter 3. Goal conflicts were evident in the gathered scenarios. These were often solved locally with routines that optimise across multiple conflicting goals, dubbed 'local ingenuity'. These routines were not necessarily at odds with the literal wording or intent of rules and regulations. However, they were not originally intended, not designed into the rules, and not explicitly included in the current documentation. Local ingenuity was either created within the existing rule base or was not transparent to the organisation outside the unit. Therefore, local ingenuity does not get tested for introducing new risks, does not become part of the knowledge base of the organisation and is not disseminated as a best practice to colleagues outside the unit.

The non-transparency of local ingenuity is explained by two paradigms regarding rule management (Hale & Borys, 2013a). The RNLAF applies the classical top-down, rational approach to rules (Hale & Borys,

2013a), also called model 1 (Dekker, 2014), whereas operators abide by a bottom-up constructivist view of rules (Hale & Borys, 2013a), also called model 2 (Dekker, 2014), as seen in the examples of local ingenuity. The model 1 paradigm focuses on a priori devised rules and procedures that incorporate the best way to do the work. Model 1 is unable to encapsulate the active complexity of a work floor. Model 2 sees rules as the dynamic, local, situated constructions of expert operators.

This study shows that depending on the model 1 paradigm is insufficient to understand how work is done. It constrains the development and transparency of local ingenuity and inhibits finding solutions for conflicting goals. As a result, local ingenuity is suboptimal as is goal realisation, reducing the overall effectiveness of the unit. Furthermore, the results of this study suggest that organisations will be more effective in solving goal conflicts and creating transparency on local ingenuity if they adopt a constructivist paradigm, if needed, combined with a classical paradigm. This includes well-defined autonomy (freedom-in-a-frame), a responsive attitude to rule modifications, operator input, and dynamic management of exceptions to meet the complexity of the operation.

Sub-question 3: Why do differences between the way work is done and prescribed endure despite their undesirability? (Chapter 4)

As seen Chapter 3, conflicting goals and the resulting local ingenuity lead to gaps between the way work is done and prescribed. These gaps are considered undesirable, yet attempts to breach persistent gaps, aligning work-as-done with work-as-prescribed, remain stymied because of a forceful balance of dynamic interests that both produce and retain the status quo, while offering all stakeholders outwardly legitimate ways to ignore its existence.

The study described in Chapter 4 shows that military top management walks a tightrope, maintaining the opportunity for differences between paper and practice, but agreeing superficially with the focus on compliance, more strongly when undesired results become officially known, as in reports of incidents and accidents. The pressure to sustain this promise of compliance leads to unsuccessful attempts to both gain discretionary space and reduce the number of rules. It makes change unlikely. This study reveals not only the self-sustaining forcefield of interests that simultaneously retain *and* deny non-compliance at almost all hierarchical levels studied, but also the routine invocation of a nondescript organisational force-to-be-feared – reminiscent of a bogeyman – in case non-compliance is discovered. The bogeyman is a symbol of a belief system that equates compliance with safety that is adhered to by politics and public and given lip service by military leaders. It is because of this belief system gaps between paper and practice endure.

Efforts to reduce the gap between paper and practice should be directed at this belief system. Pressuring employees into compliance feeds the bogeyman fear, until the fear of freedom to exercise professional autonomy stifles the entire organisation. This study shows how the achievement of 5GAF behaviour and culture is hampered due to the enduring differences between paper and practice. What could provide a solution is a change in doctrine so that rules as principles become rules as guidelines, embracing the ideals of the 5GAF.

Sub-question 4: What complexities arise when attempting to foster a restorative just culture in a mostly retributive organisation? (Chapter 5)

Chapter 5 presented the case of a safety standdown in the RNLAF to examine the complexities of fostering a restorative just culture in practice. The study found both a retributive and a restorative response. The restorative response – thus a restorative just culture – is aimed at attaining a greater level of safety in the RNLAF. The findings revealed four factors, complexities, vital to a restorative response: 1) a willingness for one to show vulnerability without being punished for this; 2) moral courage and a fair attitude, even if it means sacrificing one's own interests; 3) a management that does not punish COs for responding restoratively; and 4) a desire to learn from the mistake instead of punishing for it. However, the findings also demonstrate how hard it is to aspire to a restorative just culture in a predominantly retributive setting and culture. Examples that fostered restorative openness were threatened by the deeply engrained response to punish in the RNLAF. This made the safety standdown a rare and exceptional example. Although it was intended to create a restorative setting, it only partially accomplished this goal and even then, only spontaneously (by chance). Therefore, this particular safety standdown cannot be seen as a restorative measure but as a restorative response in an overall retributive setting. The literature suggests that restorative responses have positive effects in both the short and long term. This was also visible in this case. Relationships and trust were restored, the reasons for making the decisions were understood, the community learned from the incidents, and gossip and calls for punishment ceased. These four complexities need to be in place to strengthen the restorative response and build the restorative just culture that will enhance the safety of the organisation.

Sub-question 5: Are micro-experiments enough to reduce the enduring gaps between work-as-imagined and work-as-done? (Chapter 6)

This chapter explored whether combining Hale and Borys's framework (2013b) with micro-experiments would reduce the gap between prescription and practice, changing the rules to better match daily operations, and whether this combined approach can break through the organisational complicity.

We found that the process of micro-experiments were technically successful in showing the effectiveness of the interventions and generating shop floor commitment and the combined approach showed that it was possible to reduce gaps between work-as-done and work-as-imagined. However, we also found that actually modifying rules was impeded by the previously described tension between compliance and adaptation and the presence of a bogeyman belief regarding non-compliance. This study confirms that Dekker (2018) and De Boer's (2021) micro-experiments are effective interventions and a good means to gain shop floor commitment, but our findings point out that this method overlooks the difficulty of actually modifying rules and breaking through organisational complicity.

Our study also shows that the combination (framework and micro-experiments) only partly ended the bogeyman belief. The bogeyman reappears the moment rules need to be modified. The technical alignment solution, proposed by Hale and Borys (2013b) and De Boer (2021) will not suffice if the tension between compliance and adaptation is unresolved, keeping the bogeyman alive. The combined approach used in this study was insufficient to break through the organisational complicity. To achieve this, it seems to require an additional focus on psychological safety, restorative justice, psychological contracts and betrayal, and moral disengagement.

This means that enhancing safety requires more than a technical solution to reduce the gap between paper and practice and even more than a paradigm change from the classical to the constructivist paradigm. It seems to require expanding the constructivist paradigm with knowledge of these topics. Future research that includes these scientific domains is recommended to understand the underlying patterns and achieve real progress in aligning paper and practice.

4. Discussion

The MoD desire to enhance safety in the Defence organisation at large, including the RNLAF, is backed by political mandate and put under pressure by public and media expectations. To enhance safety the MoD focuses on the topics of safety culture, just culture and compliance and adaptation.

4.1 The focus on safety culture

As described in the Introduction, the measures the MoD is taking to achieve the desired safety culture are all based on the Safety I/model 1 paradigm given that all focus on changing (controlling) employee behaviour (Ministerie van Defensie, 2018). However, the asymptote, as articulated by Amalberti (2001) and Dekker and Pitzer (2016), shows that as a certain level of safety is reached, the effect of safety I measures decreases.

However, since improving safety culture is a recurring theme within the MoD (Ministerie van Defensie, 2018), the first study investigated whether a good safety culture actually means something.

Because previously investigated incidents led to the conclusion that the MoD lacked a good safety culture, we looked at the predictive value of safety culture, because according to the investigation reports this would prevent future accidents. Although Antonsen (2009) had already shown that safety culture has no predictive value, a more detailed safety culture assessment (SCA) might show otherwise. This proved not to be the case. Major accidents such as the Deepwater Horizon accident (Deepwater Horizon Study Group, 2011) show that demonstrated safety, such as found by an SCA, does not provide insight into the likelihood of a major accident (Rae & Provan, 2019).

Nonetheless, because measurements of safety culture can be used to demonstrate safety to external parties, for instance to politicians, it is remains difficult to stop doing these assessments (Rae & Provan, 2019).

Despite a lack of consensus, safety culture seems to be about what people do to bring about safety. However, within the RNLAF, it is not known what people actually do to achieve safety, as the focus is mainly on undesirable outcomes, that is, the absence of safety. It does seem that operators in particular create safety, as an independent committee has concluded that the safety culture amongst operators is better than expected (Visitatiecommissie Defensie en Veiligheid, 2021). But it remains unclear what this safety culture is. If this is the case, then it does not make sense to conduct an SCA based on the Safety I paradigm, with its preconceived propositions that may not fully reflect practice, based on a concept that is still unclear and that is known to have limited predictive value for actual safety outcomes

(Antonsen, 2009) (Boskeljon-Horst, et al., 2022). If mapping safety is desired, then an accurate perception of practice, defining exactly how operators create safety, has to emerge first.

4.2 An alternative focus

To provide an alternative focus, one that starts from the Safety II (Hollnagel, 2014) / Safety Differently (Dekker, 2014) paradigm, four other studies were conducted in order to discover how to describe and enhance safety and effectiveness. These four studies show that operators create safety by applying local ingenuity to solve conflicts. Local ingenuity is based on a constructivist approach to rules, which focuses on translation and adaptation before applying the rules (Hale & Borys, 2013a). The operators were confronted with rules that hampered their practice. The local ingenuity found is the result of translation and adaptation. These findings fit the findings of Tucker (2004) with nurses working around operational failures.

For instance, as Chapter 3 mentions, operators are obliged to take certain rules and procedures physically with them. However, the digital version provided had a negative impact on safety. So, the operators adapted the official version into something that matched their practice. There creative solution was a master pdf file of all the necessary rules and procedures that greatly enhanced safety. This example also shows that local ingenuity enhanced the effectiveness of the squadron, because now it takes less time to find the right rule or procedure and that enhances the effective use of available flight time (Boskeljon-Horst, et al., 2022). Mikkers et al. (2013) found the same process occurring in the case of maritime pilots guiding ships from shore when boarding ships is impossible due to adverse weather conditions. Here the pilots adapted safety margins based on the meta-knowledge they had gained from experience.

The top-down classical approach to rules (Hale & Borys, 2013a) hinders operators from creating or enhancing both safety and effectiveness because it limits local ingenuity to caveats in the rules or limits the visibility of local ingenuity to others outside the operational unit (including the rest of the military organisation) (Boskeljon-Horst, Boer, et al., 2022). This is compounded by the public and political commitment to compliance made by the MoD, and therefore the RNLAF. This has led to the belief that compliance equals safety, which is reminiscent of a bogeyman (Boskeljon-Horst, De Boer, et al., 2023). This belief is not unique to the RNLAF or the Dutch MoD. Research shows that reduced government regulatory oversight resulting from neoliberalism has resulted in an increasing number of rules and procedures and accompanying burden of compliance in the entire Health & Safety area (Dekker, 2022).

Creating or enhancing safety is further hampered by the mostly retributive climate found in the hierarchical setting of the RNLAF (Boskeljon-Horst, et al., 2023). As a result of this retributive atmosphere, individual operators and commanders need to choose between speaking up, changing and adapting, or pursuing their career. The perception that choices like these are realistic, supported by widely known examples, stifles the adaptive attitude that operators need to create, develop and apply their local ingenuity. Tucker and Edmondson (2002) suggest that perceived psychological safety has a negative relationship with workarounds, because psychological safety results in open discussion of the encountered concerns. Halbesleben and Rathert (2008) have confirmed this negative relationship between psychological safety and workarounds.

Nevertheless, even if the local ingenuity fits in with the caveats of rules or remains invisible outside the locality, applying this ingenuity leads to enduring gaps between work-as-done and work-as-imagined (Boskeljon-Horst, et al., 2023) (Boskeljon-Horst et al., n.d.). As Hale and Borys state (2013a), these gaps can be positive if they come from using expertise to translate or adapt the rules to the local circumstances and context. However, enduring gaps means that the rules and procedures have a long-lasting hampering effect, which apparently the operators and/or the organisation are unable to address. This negative effect is the result of the top-down classical approach to rules and the public and political commitment to compliance (Boskeljon-Horst, et al., 2023).

An approach that combines the Hale and Borys framework (2013b) with micro-experiments (Dekker, 2018) can help reduce the gaps between work-as-done and work-as-imagined. But this approach is not enough to ensure a permanent effect, nor can it counter the classical approach to rules, reduce the compliance promise or lessen the retributive climate. Previous studies have found that the inability of this approach to break through organisational complicity is due to these three factors. Tresfon et al. (2022), Clay-Williams et al. (2015) and Damen et al. (2018) used the Functional Analysis Resonance Method (FRAM) (Hollnagel, 2018) to align work-as-imagined with work-as-done. In all three studies the FRAM method provided insight into the difference between work-as-done and work-as-imagined for a certain procedure. Tresfon et al (2022) focused on a hospital protocol for which the mandate resides with the target organisation. The FRAM results that prompted changes to the procedure resemble the findings of the third case, discussed in Chapter 6, based on the combined framework and micro-experiments approach. Clay-Williams et al. (2015) focused on two new, unimplemented procedures. Their FRAM results also led to changes made to these procedures. These findings lead one to question if it is easier to change procedures that have not been implemented yet and to which management is not yet committed. Damen et al. (2018) used FRAM to show improvement opportunities that were not implemented in this study.

The findings of the study in Chapter 6 show that to effectively reduce gaps between work-as-done and work-as-imagined to enhance safety and effectiveness by aligning paper with the complexity of practice requires changing the beliefs about safety within the organisation (Rae & Provan, 2019), i.e. the belief that compliance equals safety. It also requires a focus on concepts such as psychological safety and restorative just culture.

4.3 Extending the model of compliance and local ingenuity

The results show a vicious cycle, reflecting the model 1/classical paradigm of rule management. Adding to the model presented in Chapter 4, Figure 1 visualises the findings of the four studies.

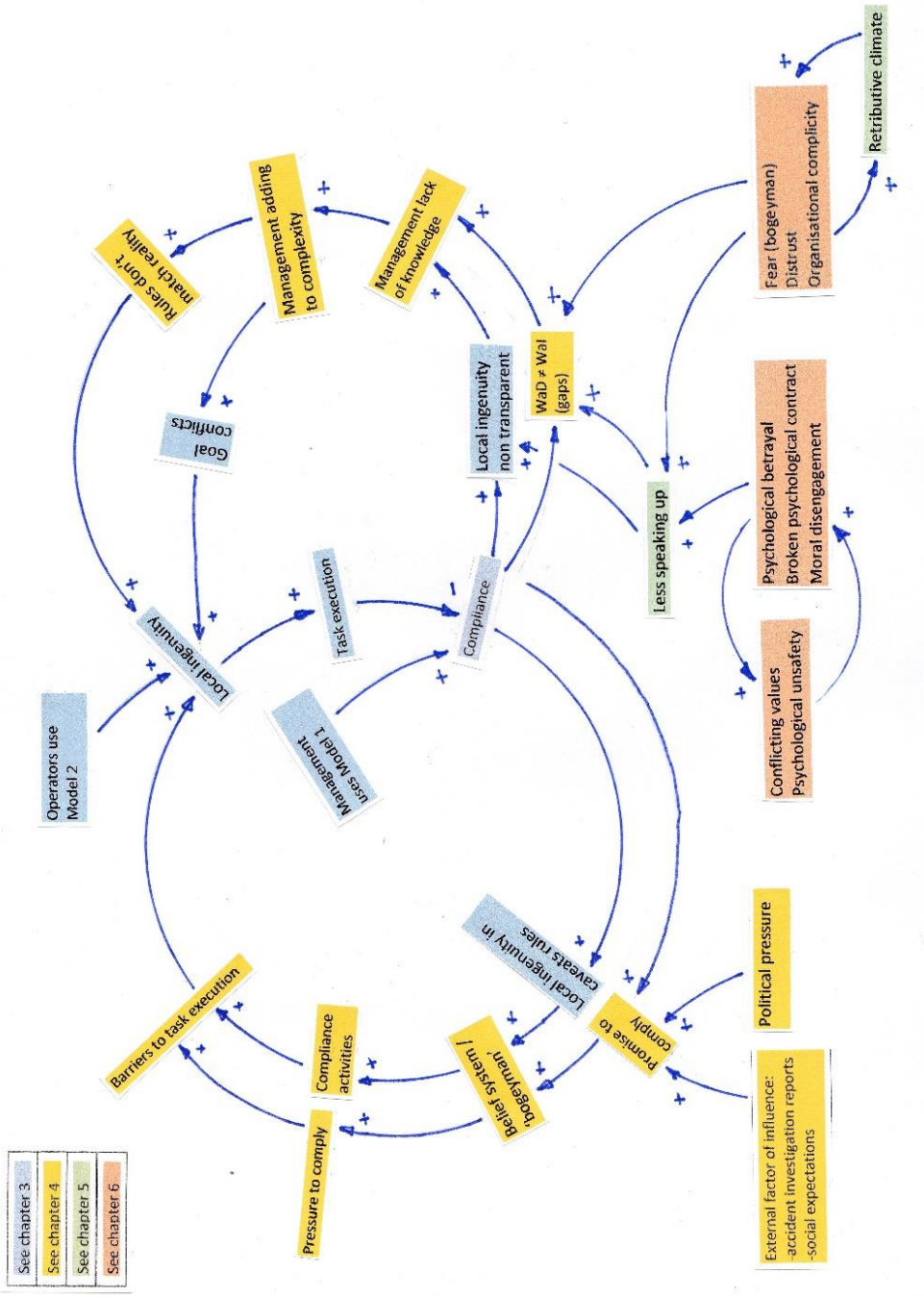

Figure 1: The creation and suppression of local ingenuity and therefore safety and effectiveness of the RNLAF.

The study presented in Chapter 3 shows that operators abide by the constructivist paradigm of rule management. To solve goal conflicts, they employ local ingenuity, which leads to enhanced goal realisation and task execution. However, the consequence is less compliance, which is important to management, as they abide to the classical paradigm of rule management. Again, reduced compliance limits local ingenuity to caveats in the rules and procedures and renders local ingenuity non-transparent to others outside the own operational unit.

The study in Chapter 4 shows that reduced compliance – influenced by the classical paradigm of rule management and combined with political pressure and external factors such as accident investigation reports and societal expectations – results in the promise to comply. Both the promise to comply and the local ingenuity limited to rule caveats lead to a belief in safety, a fear system reminiscent of a bogeyman: bad things happen if the organisation is noncompliant. This enhances both the pressure to comply and compliance activities, which in turn result in barriers to task execution. The solution, creating a full circle, is the local ingenuity found in the research described in Chapter 3. Meanwhile, however, the local ingenuity that enhances task execution and reduces compliance produces differences (gaps) between the way work is done and the way it is imagined. Management is unaware of these differences because they are unaware of (non-transparent) local ingenuity. Combined with the classical paradigm of rule management and the public/political promise to be compliant, this unawareness only adds to the complexity of daily operations when management creates new rules, especially after incidents and accidents. Often these rules do not match the reality, which creates the necessity for local ingenuity, again creating a full circle.

Finally, the studies presented in Chapters 5 and 6 show that the RNLAF's dominant retributive climate on the one hand and the fear of noncompliance (bogeyman), a lack of trust and organisational complicity on the other hand reinforce each other, resulting in inhibiting operators from speaking up and positively influences the differences between paper and practice. A reinforcing cycle of conflicting values, psychological unsafety, psychological betrayal, broken psychological contract and moral disengagement also results in an inhibition to speak up, which subsequently compounds the local ingenuity being hidden (non-transparent) and the gaps between paper and practice.

4.4 The answer to the research question

Describing and enhancing both safety and effectiveness in the RNLAF depends on understanding and enhancing local ingenuity and making instances of ingenuity transparent outside the unit in which it is applied. Consequently, this requires a change in the belief system, embracing the constructivist paradigm of rule management and a restorative just culture.

4.5 Scientific support for the answer to the research question

The literature in two scientific domains – on complex systems and on safety as an adaptive capacity – supports the answer to the research question.

4.5.1 Complex systems

Complexity thinking is a specific way of thinking about the world. After Frederick Taylor introduced the notion of scientific management in 1911, predictability, stability and control became the focus of both organisations and management (Taylor, 1911). With increasing globalisation, advancing technology and organisations becoming bigger and bigger, today's organisations exist in networks of interdependencies and face pressures to become ever faster, better and cheaper. Complexity thinking shows new ways to think about organisations and leadership and provides a means to address the current challenges and opportunities (Snowden & Boone, 2007).

Systems, including organisations, distinguish between simple, complicated, chaotic and complex systems (Snowden, 2002). This classification can be used to differentiate the strategies organisations may use as well as to decide on the appropriate organisational structure (Van Beurden et al., 2013). It helps leaders understand the context they operate in and what is expected of them (Snowden & Boone, 2007). In simple systems, cause-and-effect relationships are self-evident and understandable for any reasonable person (Snowden, 2005). The system is stable (Snowden & Boone, 2007). The manuals and standard procedures in simple systems give predictable outcomes and recognise 'best practices' (Van Beurden et al., 2013). Behaviour is predictable (Snowden, 2002). Complicated systems still have stable cause-and-effect relationships, but these are not self-evident. It requires analysis and expert knowledge to understand them (Snowden, 2002, 2005). Instead of best practices, these systems are characterised by good practices, indicating that there are multiple acceptable options (Van Beurden et al., 2013). Both simple and complicated systems consist of linear processes, in which following rules and fact-based problem-solving are the logical course of action (Snowden & Boone, 2007). The known relationships among aspects in simple and complicated systems indicate order (Nachbagauer, 2021).

There are no visible cause-and-effect relationships in chaotic systems, for example, in an unexpected climatic catastrophe (Van Beurden et al., 2013). The chaotic system is in a state of turbulence because all connections between its elements are broken (Snowden, 2002). Finally, in complex systems the relationships between the elements can only be seen in retrospect (Snowden, 2005). The complex system is dynamic, involving a large number of continuously changing interacting elements (Snowden, 2002; Snowden & Boone, 2007). The relationships are non-linear and the patterns unpredictable. Top-down procedures do not bring order to these patterns, only create new ones (Van Beurden et al., 2013). Both the chaotic and complex systems are non-linear and unordered, which make adaptation necessary to respond to the emerging patterns (Snowden & Boone, 2007). In both these systems the whole is more than the sum of its parts (Nachbagauer, 2021).

The simple, complicated and complex systems are all visible in the RNLAF. A simple system is, for instance, booking a company car. An employee must submit a request that states the date and time of intended use so that at the specified time the employee can unlock the car electronically with his credentials. The various aircraft the RNLAF operates are examples of complicated systems. They consist of many parts, and it takes an expert to understand their workings. Flying aircraft in changing environments, in different circumstances and with conflicting goals constitutes a complex system.

Because complexity is increasing, organisations and management need to find adaptive ways to manage complexity. For instance, the ability to adapt to the changes and inherent surprises in complex systems (Woods, 2015). Working according to plan, rules and procedures is not sufficient to deal with the dynamic nature of complex systems (Woods, 2019). Initiative is crucial in adapting to complex systems.

Initiative encompasses the ability to adapt when plans no longer match practice and the willingness to adapt planned work to enhance goal realisation (Woods, 2019). The local ingenuity found is the result of operators adapting to the complexity of the local contexts in which predictability, stability and control do not apply.

4.5.2 Safety as an adaptive capacity

Because complex systems require adaption, the bottom-up constructivist approach to rule management (Hale & Borys, 2013a) and the Safety II (Hollnagel, 2014) / Safety Differently approach (Dekker, 2014) provide a better fit than the top-down classical approach or Safety I approach. Seen from these paradigms (constructivist / Safety II), safety means the presence of adaptive capacity to make things go right in dynamic, surprising and uncertain situations. Many organisations put far more into what goes right than what goes wrong, thus providing a strong argument for focusing on why things go right (Hollnagel, 2014). Trying to understand safety by focusing on what goes wrong, like accidents and incidents, is like trying to understand how to have a happy, life-long marriage by studying divorce (Vos, 2018) (Dekker & Tooma, 2022).

Dekker and Tooma (2022) created a capacity index as an alternative to Safety I measures, such as incident frequency rate or time lost because of injury. These Safety I measures do not provide a good indication of the safety of an organisation because they only focus on negative outcomes and even then do not provide insight into the conditions or context of these negative outcomes.

The alternative capacity index provides six demonstrable and measurable capacities that provide insight into the safety of an organisation. They are:

1. the capacity to acquire and maintain safety knowledge;
2. the capacity to understand the nature of operations and its risks;
3. the capacity to adequately resource for safety;
4. the capacity to respond to risks and unsafe events;
5. the capacity to demonstrate engagement and compliance;
6. and the capacity for assurance.

Five of these six capacities provide support for the research described in this thesis:

Capacity 2 means having insight into the work-as-done, including adaptation, of the people closest to the actual work environment (Carvalho et al., 2009) (Havinga et al., 2018). Insight into local ingenuity, as described in Chapter 3, shows how work is done and how risks are mitigated. What is required for this capacity is a climate which permits speaking up and discussing improvements to work (Dekker & Tooma, 2022). As described in Chapters 3 to 6, because speaking up is discouraged in the RNLAF, local ingenuity and work-as-done are unknown to management. Not having the right insights – into how work gets done, how success is achieved and what trade-offs have to be made as a result of conflicting goals – is a predictor of worker fatalities (Dekker & Tooma, 2022).

Capacity 3 focuses on goal conflicts, as safety is hardly ever the most important goal among the multiple goals operators are expected to achieve simultaneously (Dekker & Tooma, 2022). Achieving goals and solving goal conflicts usually means having to make trade-offs (Hollnagel, 2009). Adequately resourcing safety means knowing and understanding these goal conflicts and trade-offs (Dekker & Tooma, 2022). The study in Chapter 3 shows that on a daily basis operators confront conflicting goals, always between

the goal of compliance on the one hand and another goal like productivity or safety on the other. The trade-offs created by the operators' local ingenuity to solve these conflicting goals are hidden from RNLAF management, which also means that no funding is allocated to solving the problems encountered. Whether the lack of funding leads to local ingenuity or the other way around is unclear. What is clear is that the safety achieved by operators is not adequately resourced.

Capacity 4 places authority as low down in the organisation as needed to respond to an event (Loukopoulos et al., 2009) and bases decisions on expertise rather than power (Farrington-Darby & Wilson, 2006). This capacity focuses on responding to the people involved in an event. Retributive responses impede openness and learning, whereas restorative responses enhance this (Dekker & Breakey, 2016). The studies in Chapter 3 to 6 reveal that the RNLAF lacks this capacity. Chapter 6 discusses the crew endurance case, in which involving the operators in decision-making could mitigate the risks of fatigue. Chapters 3, 4 and 5 show how the retributive climate stifles operators from speaking up and sharing innovative solutions to get the job done safely.

Capacity 5 concentrates on rules that have an impact on safety (Dekker & Tooma, 2022). Many rules have no relationship with safety and are frustrating for operators trying to get the job done (Dekker, 2018). Because an organisation needs to comply with such a large number of rules, the focus on safety is lost, and needs to be restored (Dekker & Tooma, 2022). The study in Chapter 4 shows that the RNLAF operators are confronted by an ever-increasing number of rules, many of which they cannot remember or bear any relationship with safety. Safety rules matter to them, but these are sometimes considered less important than, for instance, the rules made to ensure the nuisance for people living near military airbases or training areas is as little as possible.

Capacity 6 looks at the dynamic nature of complex systems. This requires the control of adaptive capacity, as opposed to the control of critical risks, human behaviour or incident numbers, to assure safety (Dekker & Tooma, 2022). As the study in Chapter 4 shows, the MoD and thus the RNLAF assure safety to politicians, public and press by making a promise of compliance, which is not achievable in the complex, dynamic, daily practice.

The five capacities of this index (Dekker & Tooma, 2022) support the findings of this thesis and confirm that they accurately describe what an organisation needs to achieve its goal of enhancing safety and effectiveness.

4.5.3 Auftragstaktik and command based warfare

Military literature also supports the findings on the positive impact of local ingenuity and the negative impact of a lack of discretionary space to execute expertise. The military equivalent of focusing more on performance goals and process rules instead of action rules (Hale & Borys, 2013a) was first called 'Auftragstaktik' [directive control], a principle developed by the Prussian Army after its defeat by Napoleon in 1806. Two causes were found for this defeat: the chaos and therefore inherent uncertainty of the battlefield and the lack of initiative in the Prussian Army due to stern discipline (Bungay, 2005). The answer was found in defining the 'why' and 'what' of a task but letting commanding officers decide on the 'how'. All hierarchical levels needed to be aware of the intentions of high command to ensure all initiatives undertaken would support the overall goal (Leussen, 2019).

Auftragstakik developed into command-based warfare. The three most important elements of command-based warfare are: clear purpose or goal (the 'why' and 'what'), freedom of action (the 'how') determined

by subordinates, and trust, both vertical and horizontal in the hierarchy. Although part of military doctrine since the 1980s, the Dutch MoD has rarely used auftragstakik in practice. Three causes are identified for this paradox (Leussen, 2019). First, political sensitivity, which moved decision-making on the use of violence to the highest level. Strict rules of engagements and other directives resulted in vanishing trust and risk-averse behaviour among operators. Second, as a result of military superiority during missions in Iraq and Afghanistan, rapid decision-making has been replaced by highly detailed planning. Third, due to long-term budget cuts military training has been reduced, resulting in less available expertise.

Although meant for the battlefield, the principles of command-based warfare also apply to daily operations. The findings of the research in this thesis strike a remarkable resemblance to the findings of a major in the Dutch army, Van Leussen (2019) on political sensitivity in a compliance climate. At a meta level Van Leussen's analysis underpins my research. It shows the Dutch MoD that people will provide the competitive advantage in a battlefield, yet this voice seems to be lost in the classical top-down approach of this organisation.

4.6 Fifth Generation Air Force

The research described in this thesis fits in well with the RNLAF transition to a Fifth Generation Air Force (5GAF), as an important means to help improve the competitive advantage of the RNLAF. For the RNLAF, the transition means embracing such concepts as ownership, continuous learning and reflection, adaptability, resilience, agility and flexibility. In a time when the opponent has access to the same technology and equipment, the RNLAF considers the air force employee key in making the difference (Defensie, n.d.). The two studies presented in Chapters 3 (Boskeljon-Horst, Boer, et al., 2022) and 4 (Boskeljon-Horst, et al., 2023) contribute to this initiative as they have shown that both enhancing local ingenuity and resolving goal conflicts as well as gaining insight into the remaining differences between the way work is done and prescribed. The studies also show that lack of transparency on local ingenuity and conflicting paradigms on compliance and adaptation hamper the achievement of 5GAF behaviour and culture. The RNLAF and its management have little idea how to achieve this competitive advantage, as internal documents show. This thesis can add to the small body of current knowledge and play an active part in facilitating the desired advantage.

4.7 What is needed to enhance RNLAF safety and effectiveness

The extended model in Figure 1 shows that safety together with effectiveness is a concept that requires an integrated approach. It also shows why change must be found in a paradigm shift from model 1 to model 2 thinking, not only specifically for rule management, but also more generally for to manageability and predictability. Model 1 thinking assumes the correctness of rules, predictability and manageability: if you follow the rules then you are safe. If you do not follow the rules, you are not safe. Rules provide control over the concept of safety. But many processes in the RNLAF fall into the realm of complexity and complex systems. A characteristic of complexity is that processes are non-linear and therefore not controllable or predictable. Model 1 thinking does not fit complexity hence a paradigm shift to a constructivist approach to rule management, model 2 thinking, is needed to put the compliance commitment in perspective and enhance both safety and effectiveness.

Furthermore, it needs a paradigm shift from a retributive to a restorative just culture. The RNLAF sees value in punishing employees for negative outcomes because this resembles and supports the political and public commitment to compliance, as Chapter 5 has shown. The idea that learning and punishment work together is strong within the RNLAF, although it has long been known these are mutually exclusive (Dekker & Tooma, 2022). The research in this theses shows, however, that the retributive climate does not support learning. In fact, it hampers learning by stifling speaking up and hiding local ingenuity.

Finally, the RNLAF needs to start living up to the statements made in writing on trusting and valuing employee expertise. So far, this expertise has only been trusted within the boundaries of the existing rule base. However, the local ingenuity found shows that operators need more space in which to use their expertise. This requires freedom-in-a-frame – rules as a framework within which operators have the freedom to use their expertise – (De Boer, 2021) or discretionary space for professional judgement (Biermann-Teuscher et al., 2022). Providing the space to exercise expertise has all sorts of benefits. For instance, Carmeli and Spreitzer (2009) found that trust and the promotion of growth ultimately increased innovative behaviour in employees. Dekker and Laursen (2007) found that when employees realised they could make a difference in their organisation, it had a positive influence on their willingness to report errors and mistakes. Dekker (2007) also found that employees are more willing to take on the responsibilities that come with their jobs when they are involved in changing and improving their work conditions. Grabowski et al. (2007) found that, among other factors, employee empowerment is significantly correlated with safety performance of the organisation. Perhaps what the RNLAF needs to enhance safety and effectiveness can best be explained by General George S. Patton: 'Never tell people how to do things. Tell them what to do, and they will surprise you with their ingenuity.'

5. Conclusions

The research question of this thesis is: How can we describe and enhance the safety and effectiveness of the RNLAF? In studying the RNLAF with its top-down classical approach to rule management, political and public commitment to compliance and focus on negative outcomes when it comes to safety, it has become clear that local ingenuity is the key first step in describing and enhancing both safety and effectiveness. The RNLAF has to know about, recognise, understand and enhance the local ingenuity of operators and make their ingenuity transparent outside the unit in which it is applied. The RNLAF has to know about and understand the conflicting goals that lead to local ingenuity and the trade-offs that need to be made to get the job done. Actively looking for goal conflicts and local ingenuity requires both changing the belief system into one that embraces the constructivist paradigm of rule management and developing a restorative just culture. These three changes will ensure the needed psychological safety that enables operators to speak up about goal conflicts, trade-offs and local ingenuity. Knowing about local ingenuity and its effects, such as goal achievement, leads to knowledge of safety and effectiveness and is therefore a way to enhance both. Knowledge of the tension in the RNLAF that comes from its classical approach to rule management and such patterns as the compliance promise, the bogeyman, the retributive just culture and lack of psychological safety all show what stands in the way of increasing safety and effectiveness and therefore shows what the RNLAF needs to do to enhance both.

The findings of this thesis my also benefit other organisations resembling the RNLAF. This includes other departments in the Defence organisation, such as the Royal Army and Royal Navy. Considering the fact that the compliance goal is the same throughout the entire Defence organisation, it is highly likely that the army and navy face the same conflicting goals that demand creative resolution. Furthermore, although this research took place within a military organisation, any other organisation that has a firm compliance policy that creates pressure on other organisational goals will no doubt find the information on local ingenuity intriguing.

The research in this thesis contributes the following to current scientific knowledge:

1. The conclusion that safety culture maturity assessments not only have limited predictive value but may in fact inversely predict safety outcomes.
2. The concept of local ingenuity, a non-normative term, which refers to routines created to solve goal conflicts that are part of operators' regular repertoire, but are relatively invisible to management.
3. Mapping the benefits, fears and denial of non-compliance has gained insight into the interests behind retaining undesired gaps between work-as-prescribed and work-as-done.
4. Four factors that need to be overcome to foster a restorative just culture in a predominantly retributive organisation are: 1) a willingness to be vulnerable without being punished for this; 2) moral courage and a fair attitude, even if it means sacrificing one's own interests; 3) management that does not punish COs for responding restoratively; and 4) a desire to learn from the mistake instead of punishing for it.
5. The conclusion that De Boer's proposal to combine a framework of rule management with micro-experiments is valuable yet insufficient to break through organisational complicity and achieve a lasting reduction in the gaps between work-as-imagined and work-as-done.

In practical terms of achievement, what this thesis contributes to the RNLAF is a ground-breaking description, previously unknown within this military organisation, of what is needed for the transition into a true Fifth Generation Air Force in which employees – RNLAF people – will provide competitive advantage: creating freedom-in-a-frame for local ingenuity to thrive.

6. Proposals for future research

Since this research was conducted with all male interviewees, it could be interesting to see if its findings on goal conflicts, the creation and execution of local ingenuity, and goal realisation would be the same in target entity with both male and female interviewees.

Further and deeper ethnographic research would be interesting to form a thorough understanding of the function and role of the 'bogeyman' in sustaining concurrent acquiescence and non-compliance.

Further research into actually reducing the gaps between work-as-imagined and work-as-done to achieve lasting change in the existing rule base is needed to add to the literature on enduring gaps.

References

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science, 37*(2), 101–118. https://doi.org/10.4324/9781315092898-7

Antonsen, S. (2009). Safety culture assessment: A mission impossible? *Journal of Contingencies and Crisis Management, 17*(4), 242–254. https://doi.org/10.1111/j.1468-5973.2009.00585.x

Biermann-Teuscher, D., Thissen, L., Horstman, K., & Meershoek, A. (2022). Safety: A Collective and Embedded Competency. An Ethnographic Study of Safety Practices at an Industrial Workplace in the Netherlands. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4094007

Boskeljon-Horst, L., Boer, R. J. De, Sillem, S., & Dekker, S. W. A. (2022). Goal Conflicts , Classical Management and Constructivism : How Operators Get Things Done. *Safety, 8*(37).

Boskeljon-Horst, L., De Boer, R. J., & Dekker, S. W. A. (n.d.). Reducing gaps between paper and practice requires more than a technical alignment. *Safety Science, under Review*.

Boskeljon-Horst, L., De Boer, R. J., Steinmetz, V., & Dekker, S. W. A. (2023). Aircrews, rules and the bogeyman: Mapping the benefits and fears of noncompliance. *Safety, 9*(1).

Boskeljon-Horst, L., Sillem, S., & Dekker, S. W. A. (2022). ' Ladder ' -based safety culture assessments inversely predict safety outcomes. *Journal of Contingencies and Crisis Management, January*, 1–20.

Boskeljon-Horst, L., Snoek, A., & Baarle, E. van. (2023). Learning fom the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force. *Safety Science, 161*.

Bungay, S. (2005). The Road to Mission Command: The Genesis of a Command Philosophy. *British Army Review, 137*(6).

Carmeli, A., & Spreitzer, G. M. (2009). Trust, connectivity, and thriving: implications for innovative behaviors at work. *Journal of Creative Behavior, 43*(3), 169–191. https://doi.org/10.1002/j.2162-6057.2009.tb01313.x

Carvalho, P. V. R. de, Gomes, J. O., Huber, G. J., & Vidal, M. C. (2009). Normal people working in normal organizations with normal equipment: System safety and cognition in a mid-air collision. *Applied Ergonomics, 40*(3), 325–340. https://doi.org/10.1016/j.apergo.2008.11.013

Clay-Williams, R., Hounsgaard, J., & Hollnagel, E. (2015). Where the rubber meets the road: Using FRAM to align work-as-imagined with work-as-done when implementing clinical guidelines. *Implementation Science, 10*(1), 1–8. https://doi.org/10.1186/s13012-015-0317-y

Damen, N. L., de Vos, M. S., Moesker, M. J., Braithwaite, J., de Lind van Wijngaarden, R. A. F., Kaplan, J., Hamming, J. F., & Clay-Williams, R. (2018). Preoperative Anticoagulation Management in Everyday Clinical Practice. *Journal of Patient Safety, 00*(00), 1.

De Boer, R. J. (2021). *Safety Leadership. A different, doable and directed approach to operational improvements*. CRC Press.

Deepwater Horizon Study Group. (2011). Final Report on the Investigation of the Macondo Well Blowout. http://ccrm.berkeley.edu/pdfs_papers/bea_pdfs/DHSGFinalReport-March2011-tag.pdf

Defensie. (n.d.). 5e generatie luchtmacht - wat houdt het in? https://www.youtube.com/watch?v=5p9KXOrJbak

Dekker, S. W. A. (2007). Criminalization of medical error: Who draws the line? *ANZ Journal of Surgery, 77*(10), 831–837. https://doi.org/10.1111/j.1445-2197.2007.04253.x

Dekker, S. W. A. (2014). *Safety Differently. Human factors for a new era.* Apple Academic Press Inc.

Dekker, S. W. A. (2018). *The Safety Anarchist*. Routledge Taylor & Francis Group.

Dekker, S. W. A. (2022). *Compliance Capitalism: How Free Markets Have Led to Unfree, Overregulated Workers*. Routledge Taylor & Francis Group.

Dekker, S. W. A., & Breakey, H. (2016). "Just culture:" Improving safety by achieving substantive, procedural and restorative justice. *Safety Science, 85*, 187–193. https://doi.org/10.1016/j.ssci.2016.01.018

Dekker, S. W. A., & Laursen, T. (2007). From Punitive Action to Confidential Reporting. *Patient Safety & Quality Healthcare, 45*(19), 1–8. http://cla.sagepub.com/cgi/doi/10.1177/0069477007310054

Dekker, S. W. A., & Pitzer, C. (2016). Examining the asymptote in safety progress: A literature Review. *International Journal of Occupational Safety and Ergonomics, 22*(1), 57–65. https://doi.org/10.1080/10803548.2015.1112104

Dekker, S. W. A., & Tooma, M. (2022). A capacity index to replace flawed incident-based metrics for worker safety. *International Labour Review, 161*(3), 375–393. https://doi.org/10.1111/ilr.12210

Farrington-Darby, T., & Wilson, J. R. (2006). The nature of expertise: A review. *Applied Ergonomics, 37*(1 SPEC. ISS.), 17–32. https://doi.org/10.1016/j.apergo.2005.09.001

Grabowski, M., Ayyalasomayajula, P., Merrick, J., Harrald, J. R., & Roberts, K. (2007). Leading indicators of safety in virtual organizations. Safety Science, 45(10), 1013–1043. https://doi.org/10.1016/j.ssci.2006.09.007

Halbesleben, J. R. B., & Rathert, C. (2008). The role of continuous quality improvement and

psychological safety in predicting work-arounds. *Health care management review 33*(2), 134–144.

Hale, A., & Borys, D. (2013a). Working to rule, or working safely? Part 1: A state of the art review. *Safety Science, 55*, 207–221. https://doi.org/10.1016/j.ssci.2012.05.011

Hale, A., & Borys, D. (2013b). Working to rule or working safely? Part 2: The management of safety rules and procedures. *Safety Science, 55*, 222–231. ttps://doi.org/10.1016/j.ssci.2012.05.013

Havinga, J., Dekker, S. W. A., & Rae, A. (2018). Everyday work investigations for safety. *Theoretical Issues in Ergonomics Science, 19*(2), 213–228. https://doi.org/10.1080/1463922X.2017.1356394

Hollnagel, E. (2009). *The ETTO principle*. Taylor & Francis Ltd.

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management*. Taylor & Francis Group.

Hollnagel, E. (2018). *Safety-II in practice Developing the resilience potentials*. Routledge Taylor & Francis Group.

Leussen, W. L. van. (2019). Opdrachtgerichte Commandovoering. *Militaire Spectator, 188*(2), 60–71. https://www.militairespectator.nl/thema/operaties-leiderschap/artikel/opdrachtgerichte-commandovoering%0Ahttps://www.militairespectator.nl/sites/default/files/teksten/bestanden/Militaire Spectator 2-2019 Van Leussen.pdf

Loukopoulos, L. D., Dismukes, K., & Barshi, I. (2009). *The multitasking myth:Handling complexity in real-world operations*. Ashgate Publishing Limited.

Mikkers, M., Henriqson, E., & Dekker, S. W. A. (2013). Managing multiple and conflicting goals in dynamic and complex situations: Exploring the practical field of maritime pilots. *Journal of Maritime Research, 9*(2), 13–18.

Ministerie van Defensie. (2018). Een veilige defensieorganisatie. Plan van aanpak te verbetering van de veiligheid binnen Defensie.

Nachbagauer, A. (2021). Managing complexity in projects: Extending the Cynefin framework. *Project Leadership and Society, 2*, 100017. https://doi.org/10.1016/j.plas.2021.100017

Rae, A., & Provan, D. (2019). Safety work versus the safety of work. *Safety Science, 111*, 119–127. https://doi.org/10.1016/j.ssci.2018.07.001

Snowden, D. (2002). Complex acts of knowing: Paradox and descriptive self-awareness. *Journal of Knowledge Management, 6*(2), 100–111. https://doi.org/10.1108/13673270210424639

Snowden, D. (2005). Strategy in the context of uncertainty. *Handbook of Business Strategy, 6*(1), 47–54. https://doi.org/10.1108/08944310510556955

Snowden, D. J., & Boone, M. E. (2007). A Leader's Framework for Decision Making. *Harvard Business Review, November*.

Taylor, F. W. (1911). *The principles of scientific management*. Harper & Brothers Publishers.

Tresfon, J., Brunsveld-Reinders, A. H., Van Valkenburg, D., Langeveld, K., & Hamming, J. (2022). Aligning work-as-imagined and work-as-done using FRAM on a hospital ward: a roadmap. *BMJ Open Quality, 11*(4), 1–9. https://doi.org/10.1136/bmjoq-2022-001992

Tucker, A. L. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management, 22*(2), 151–169. https://doi.org/10.1016/j.jom.2003.12.006

Tucker, A. L., Edmondson, A. C., & Spear, S. (2002). When problem solving prevents organizational learning. *Journal of Organizational Change Management, 15*(2), 122–137. https://doi.org/10.1108/09534810210423008

Van Beurden, E. K., Kia, A. M., Zask, A., Dietrich, U., & Rose, L. (2013). Making sense in a complex landscape: How the cynefin framework from complex adaptive systems theory can inform health promotion practice. *Health Promotion International, 28*(1), 73–83. https://doi.org/10.1093/heapro/dar089

Visitatiecommissie Defensie en Veiligheid. (2021). Ruimte voor veiligheid. Eindrapport 2021.

Vos, M. S. de. (2018). *Healthcare improvement based on learning from adverse outcomes*. Doctoral Thesis, Leiden University Medical Center.

Woods, D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering and System Safety, 141*(September), 5–9. https://doi.org/10.1016/j.ress.2015.03.018

Woods, D. (2019). Essentials of resilience, revisited. *Handbook on Resilience of Socio-Technical Systems, January*, 52–65. https://doi.org/10.4337/9781786439376.00009

List of Figures and Tables

List of Figures

List of Tables

Page left intentionally blank

About the author

Curriculum Vitae Leonie Boskeljon-Horst

Leonie Boskeljon-Horst, born the 7th of December 1977 in Eindhoven, the Netherlands, studies social and organisational psychology at the University of Leiden (1996-2000).

After obtaining her master degree, she was employed by the RNLAF and subsequently followed the RNLAF post master study to become an aviation psychologist (2002-2006).

She worked as an aviation psychologist for over 17 years. During these years she was responsible for human performance and safety within military aviation (mostly helicopter crews) by:

- conducting hundreds of safety training programs, such as Crew Resource Management, Military Non-Technical Skills Training, Training Skills for Trainers and Instructors and Human Performance for the Elementary Military Aviation Training;
- coaching pilots and crewmembers in case of performance issues, mental/emotional issues and as a form of care in case of deployments;
- providing (unsolicited) human performance and safety advice to squadron and airbase management;
- participating in investigation committees after an incident or accident, specifically to investigate the 'human factor' of the event.

In 2018 she became Head of Occurrence Investigations at the RNLAF Central Safety Department, specifically to innovate the investigation process. In the years 2018-2023 she introduced the Safety II paradigm to the RNLAF, drastically changed the way occurrence investigations are done, participated in a Defence-wide Just Culture and Psychological Safety action research and is frequently invited to speak at Defence and/or RNLAF symposia, safety days and training programs.

She started the research described in this dissertation in 2021, with Prof. dr. ir. Pieter van Gelder and Prof. dr. Sidney Dekker as promotors and Dr. Simone Sillem as copromotor. The research focused on the Defence Helicopter Command, the command she was working with as an aviation psychologist.

Education

2017	Human Factors and Safety
	Hogeschool van Amsterdam
2005	Military education to become a Major
	Instituut Defensie Leergangen
2003 - 2005	Courses part of postmaster education Aviation Psychologist
	University of Southern California
2000 – 2001	Military academy to become an officer
	Royal Military Academy
1996 – 2000	Social and Organisational psychology

	University of Leiden
1990 – 1996	High school
	Oranje Nassau College, Zoetermeer

Publications

Boskeljon-Horst, L., Sillem, S., & Dekker, S.W.A. (2022). 'Ladder'-based safety culture assessments inversely predict safety outcomes. *Journal of Contingencies and Crisis Management, January* (1-20). Doi: 10.1111/1468-5973.12445

Boskeljon-Horst, L., De Boer, R.J., Sillem, S., & Dekker, S.W.A. (2022). Goal Conflicts, Classical Management and Constructivism: How Operators Get Things Done. *Safety, 8* (37). Doi: 10.3390/safety8020037

Boskeljon-Horst, L., De Boer, R.J., & Boland, E.J. (2022). Lokale vindingrijkheid bij de Koninklijke Luchtmacht. *Militaire Spectator 191* (10).

Boskeljon-Horst, L. , & Alta, M. (2023). Voorvallenonderzoek contraproductief? De negatieve impact op sociale veiligheid en lerend vermogen. *Militaire Spectator, 192* (3).

Boskeljon-Horst, L., De Boer, R.J., Steinmetz, V., & Dekker, S.W.A. (2023). Aircrews, Rules and the Bogeyman: Mapping the Benefits and Fears of Noncompliance. *Safety 9* (1). Doi: 10.3390/safety9010015

Boskeljon-Horst, L., Snoek, A., & Van Baarle, E. (2023). Learning from the complexities of fostering a restorative just culture in practice within the Royal Netherlands Air Force. *Safety Science 161* (May). Doi: 10.1016/j.ssci.2023.106074

Boskeljon-Horst, L., De Boer, R.J., & Dekker, S.W.A. (n.d.). Reducing gaps between paper and practice requires more than a technical alignment. *Safety Science, under Review.*