Joint Meeting

15th Conference on Naturalistic Decision Making

9th Symposium on Resilience Engineering

Bouncing forward from global crises and challenges

NDM, REA, FONCSI, ICSI

21-24th June 2021

Toulouse, France
Columbus, Ohio
and around the world online

Resilience Engineering Association

Naturalistic Decision Making

ICSI
Institut pour une culture de sécurité industrielle

FONCSI
Fondation pour une culture de sécurité industrielle

ENAC
Ecole Nationale de l'Aviation Civile

Preface

Vision - An organic programme that evolves with participants' contributions

The programme included content for managers or operational personnel from diverse businesses, government and communities to look at the agenda and see benefit to their organisation. The academics benefitted for their research by attending and presenting their latest work. Our speakers and contributors offered presentations and interactive sessions. These speakers and contributors represent industry, governments, communities and academia.

By 17th June 2021, we had 337 Participants from 30 countries around the globe. There were 76 submissions, and following the review process, the Programme Committee decided to accept 57.

Themes

Global crises and their reverberations, challenges and increased digital transformation are calls for society to adapt and bounce forward. Adaptation needs are everywhere, at systemic, organisational, and societal levels, as well as at individual and community levels. It is urgent to rethink and guide innovative forms of adaptation. Overall, the event explored how RE and NDM are distinct and can complement each other as a rich source of inspiration for such adaptations.

The event covered the following themes:

- New thinking and ideas; exploring complementarities and differences between NDM and RE, Safety II and Safety Differently
- Societal resilience: contributions from RE and NDM, is there a need for other disciplines
- Organisational resilience: how can organisations strengthen their resilient performance
- Challenges and opportunities associated with emergent technologies and collaborative work
- Industry sharing experience and gap between theory and practice
- Everyday work, decision making and performance improvements
- Training and learning across individuals, teams and organisations;
- COVID, remote work and other forms of change: look at challenging situations combining science and practical experience, what can we learn to move forward? New capabilities have been discovered and enabled. How can institutions leverage this crisis to sustain key adaptive capabilities beyond this pandemic?
- Cases studies where systems are challenged to deliver services examples from diverse contexts and domains e.g. software services, energy, mobility, communications, finance, government

A different event? YES!

- We organised diverse nodes for registered attendees to have the opportunity to meet in person. There was a central node in Toulouse, France and one in Columbus, Ohio. On June 23, ICSI- FONCSI initiatives were held in Toulouse, with a strong in-person attendance of French industry representatives. This was complemented with showcases around the world, namely from Scandinavia with contributors from different Norwegian and Swedish locations, from Australia and Brazil with contributors meeting in-person locally, or joining online from different cities and regional areas.

- Sessions were built around interactions, knowledge and experience sharing and lively discussions, offering the opportunity to bringing forward recent industry and academic work, new ideas, as well as emerging problems, industry needs and foreseeable challenges.

Not covered - out of the scope

Any discussion relating to solely Safety I or Risk management in isolation were strongly discouraged. All speakers were challenged to be open to explore diverse perspectives, as opposed to generating niche discussions. Efforts were also devoted in avoiding any analyses of big accidents and critical incidents that would not provide useful insights on everyday challenges and adaptation.

After the event

Results were recorded, videos and documents are being made available to help attendees to bounce forward. Speakers are invited to pursue with their thoughts based on their presentation/panel after the conclusion of the Symposium.

We are very excited about the achievements of this event and look forward to hear from participants on how progress, knowledge, science, practice, and friendships are maintained.

June 2021 The Programme Chairs
Ivonne Herrera
Jan Maarten Schraagen

List of authors

Wordcloud from keywords submitted

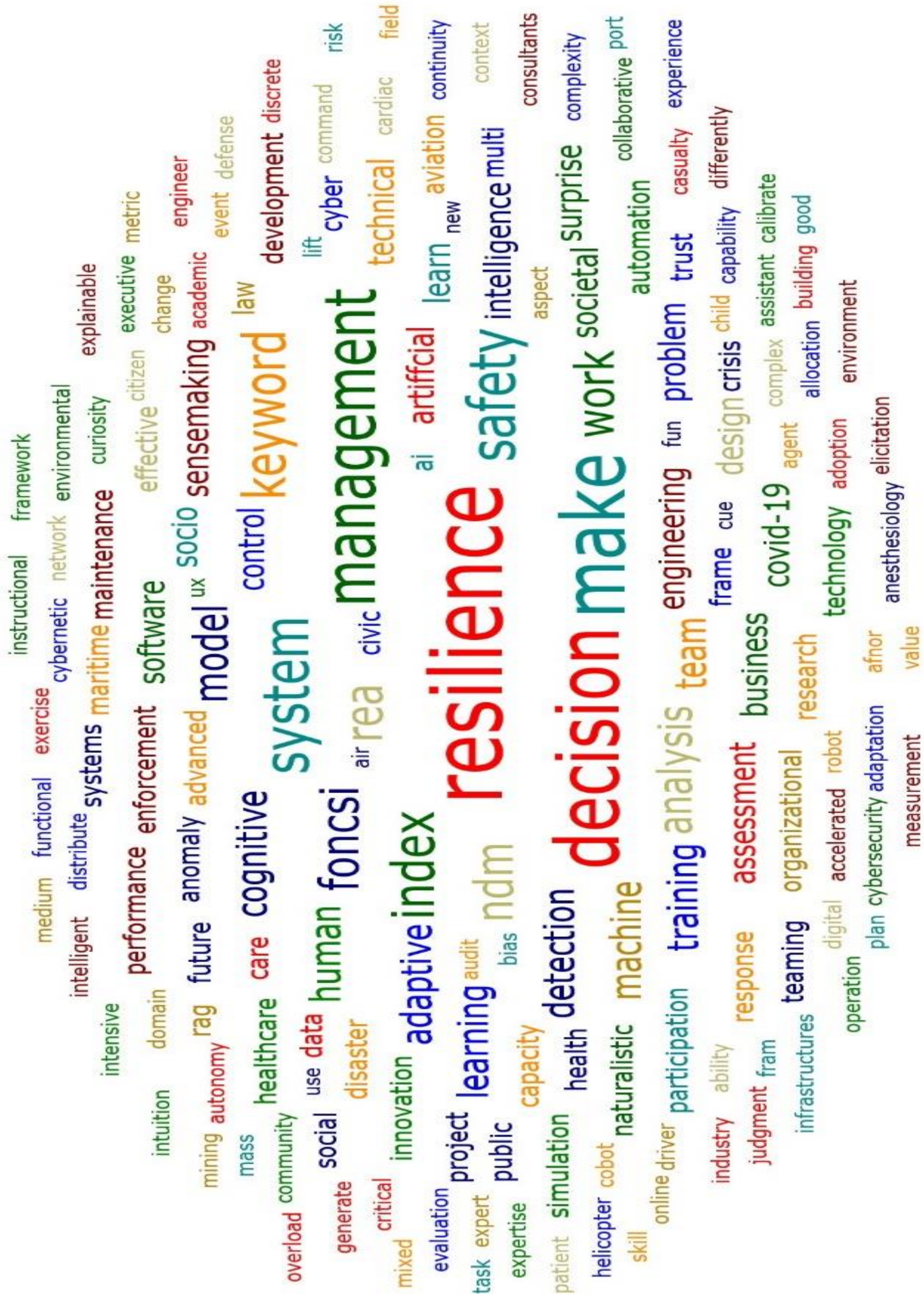


Table of contents

Highlights of recent work

Highlights of Recent Work
Interdependency Analysis for Collaborative Robot Applications Through FRAM Analysis

Arie Adriaensen[a, b], Liliane Pintelon[b], Riccardo Patriarca[a]
[a] *Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Italy*
b Centre for Industrial Management, KU Leuven University, 3001 Leuven, Belgium

- This project takes interdependency relations in terms of observability, predictability, and directability (OPD), derived from Coactive Design Theory as the basic principles to assess human-robot interaction safety in collaborative robot (cobot) applications.
- The performance variability potential from OPD requirements will be assessed in a multi-scale modelling framework for human-robot with a Joint Cognitive Systems (JCS) unit of analysis.

1. Summarize New Results

Although collaborative robots (cobot) applications are increasingly utilised to physically collaborate in real-time with human operators, current cobot safety mainly focuses on a techno-centric perspective (Guiochet, Machin, & Waeselynck, 2017) in terms of physical separation and managing the net result of kinetic energy in the cobot system. Human Factors (HF) research or its subcategory Cognitive Systems Engineering (CSE) are largely absent in cobot safety (Kadir, Broberg, & Conceição, 2019). To complement the current techno-centric approach, this project introduces a socio-technical safety and resilience analysis perspective for human-robot interaction by applying systemic safety analysis methods, with a specific focus on OPD requirements for JCS in relation to cobot operations. The scope of cobots in our project is wider than the typical industrial cobot definition and encompasses other robot operations, as long as they are defined by the potential for foreseeable or intentional physical contact between robots and humans, not necessarily restricted to system operators.

The project sets out a generic cobot safety framework that is based on a FRAM analysis of the work system and will follow the three first steps of a FRAM analysis (Hollnagel, 2012): (i) identification of functions in a given work system; (ii) identification of variability, and; (iii) aggregation of variability. The fourth and last step, being the management of variability will be omitted, but the results from the analysis will typically lead to a design reiteration of the work system, which has the same ultimate purpose of managing variability. FRAM has previously been applied to JCS analysis in which agents can be defined as any human or technical system actor or medium. (Adriaensen, Patriarca, Smoker, & Bergström, 2017)

The traditional FRAM model, based on the identification of functions will be supported by a number of additional function labels: (i) the functions will be allocated to the agents that perform the function in the work system, e.g. robot, operator, supervisor; (ii) thereafter, the functions that belong to the same functional cluster from a bottom-up grouping of clusters will be assigned to a number of JCS subunits, e.g. functions like ‘navigation’, ‘picking’, ‘releasing’, but also implicit functions like ‘separation behaviour interpretation’ or other functional implicit functional clusters that emerge from the FRAM analysis etc. Such subunits that consist of multiple individual functions will be considered as JCS agents.

The second and third step, being the identification and aggregation of variability in terms of assigning aspects and their resonance between functions will be mapped on a RAM matrix representation (Patriarca, Del Pinto, Di Gravio, & Costantino, 2018) of the traditional FRAM model. The RAM matrix will be designed in order to interpret the output variability from a multi-scale model, where both inter and intra-level resonance between agents on the one hand, and JCS subunits on the other hand can be assessed. The assessment will examine to which extent pre-defined rules or mechanisms correspond to the OPD principles in cobot systems as the essential interaction requirements to complement the physical safety aspect of human-robot separation and kinetic energy management. Observability can be defined as the capability of mutually predicting agent status; predictability as the system ability to rely on another agent’s actions while considering one’s own agent’s actions, and; directability as the ability to influence the behaviour of others as well as being influenced by others. Together these principles can be seen as the most essential interaction requirements that can be generically applied to any human-robot interaction, as the essence of joint activity is about interaction and negotiation, borrowed from Interdependency Analysis in Co-Active design (Johnson, Bradshaw, & Feltovich, 2018).

2. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Developers, integrators and users of industrial, medical and other cobot applications will benefit directly from the published results. Additionally, other Joint Cognitive work Systems will benefit from the inter- and intra-agent analysis framework as the RAM representation can be tailored to other needs.

3. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The results should push for a socio-technical complement in robot safety, which is currently governed by a techno-centric safety perspective. Developers, integrators and users will have the chance to learn about theoretical and practical implications in relation to systemic safety methods applied to the specific challenges of cobot design, cobot task analysis and workplace integration. An interdependency analysis has also been proposed to counter the substitution fallacy in functional allocation. Instead of substituting one agent for another, such as in the 'Men-are-better-at/Machines-are-better-at' (MABA-MABA) approach (Fitts, 1951), or comparing two different levels of automation through empirical evaluation, such as in the LoA approach (Parasuraman, Sheridan, & Wickens, 2000), interdependency should shape automation by taking into account the joint action potential (Johnson et al., 2018).

4. THE WIDER SCOPE OR RELEVANCE OF THE WORK

These activities were initially developed within a PhD research project but are expected to contribute to literature via publications about the generic framework and applied to case studies in the domain of industrial and medical collaborative robots.

5. REFERENCES

Adriaensen, A., Patriarca, R., Smoker, A., & Bergström, J. (2017). Can artefacts be analyzed as an agent by itself – Yes or No: what does Hutchins 'how does a cockpit remember its speeds' tell us. *7th REA Symposium*. Liège, Belgium.

Fitts, P. M. (1951). *Human engineering for an effective air-navigation and traffic-control system*. https://doi.org/10.1016/j.na.2009.04.070

Guiochet, J., Machin, M., & Waeselynck, H. (2017). Safety-critical advanced robots: A survey. *Robotics and Autonomous Systems*, *94*, 43–52. https://doi.org/10.1016/j.robot.2017.04.004

Hollnagel, E. (2012). *FRAM: The functional resonance analysis method: Modelling complex socio-technical systems*. https://doi.org/10.1017/CBO9781107415324.004

Johnson, M., Bradshaw, J. M., & Feltovich, P. J. (2018). Tomorrow's Human–Machine Design Tools: From Levels of Automation to Interdependencies. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 77–82. https://doi.org/10.1177/1555343417736462

Kadir, B. A., Broberg, O., & Conceição, C. S. da. (2019). Current research and future perspectives on human factors and ergonomics in Industry 4.0. *Computers and Industrial Engineering*, *137*(July), 106004. https://doi.org/10.1016/j.cie.2019.106004

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation - Systems, Man and Cybernetics, Part A, IEEE Transactions on. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 1–12. https://doi.org/10.1109/3468.844354

Patriarca, R., Del Pinto, G., Di Gravio, G., & Costantino, F. (2018). FRAM for Systemic Accident Analysis: A Matrix Representation of Functional Resonance. *International Journal of Reliability, Quality and Safety Engineering*, *25*(1). https://doi.org/10.1142/S0218539318500018

Highlights of Recent Work

Leveraging Mixed Methods Research Methodology for Resilience Engineering: A Case Study at American Airlines

Christine JEFFERIES[a]
[a]*Department of Integrated Systems Engineering, The Ohio State University*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

Research methodology and specific research methods for Resilience Engineering (RE) research remains sparse within the published literature. We present a mixed methods approach to data collection and analysis to support a new Safety-II/RE-informed data stream within a commercial airline's safety management system.

2. Summarize New results

American Airlines (AA), based in Fort Worth, Texas, USA, began to incorporate Safety-II and Resilience Engineering (RE) principles into their safety management system (SMS) in August 2018. The Learning and Improvement Team (LIT) chose to base their approach on the RAG model. After transforming RAG's model of four abilities into a bespoke conceptual framework for modeling resilient performance, LIT further granulated their four "potentials" by extrapolating observable flightdeck behaviors, or "proficiencies". Together with researchers from the Cognitive Systems Engineering Laboratory (C/S/E/L) at The Ohio State University, LIT designed a multi-stage mixed methods approach to collect observational and interview data on the flight deck and during off-duty hours. The research questions were: can a new Safety-II/RE-informed data stream add value to the existing SMS, and can a Safety-II/RE-informed data collection tool detect resilient performance among commercial airline pilots?

We present the evolution of LIT's research methodology to its current state, our multi-stage mixed methods data collection and analysis strategy, and a reflective critique of this overall approach:

Methodology

- Largely informed by Dr. Hollnagel's work in Safety-II, LIT seeks to learn from what goes right (or goes well, or goes as desired).
- Poly-categorical ontological refinement should continue until all stakeholders are satisfied with the *decision-driven utility* of each category.
- Data collection should follow a pragmatic epistemological approach: By training pilots to observe front-line work through a Safety-II/RE-informed lens, pilot-observers should be able to detect actions and behaviors that lead to resilient performance of pilots at work.

Mixed Methods

- In the first stage, an exploratory sequential design used qualitative interview data from check airmen to design a flight deck observation instrument to collect qualitative and quantitative data on pilot actions and behaviors during flights.

Qualitative flight deck data was transformed into quantitative prior to analysis as well as analysed on its own. The data collected with this instrument was analysed and the instrument was iteratively refined in tandem with a bespoke theoretical model and visual representation thereof. Inter-coder reliability assessments promoted cohesion and confidence among observers.
- In the second stage, a convergent design used off-duty pilot interviews together with jump seat observations completed by LIT observers using the observation instrument. Qualitative flight deck data were compared to interview data to establish concordance and were interpreted jointly using a grounded theory approach. Quantitative and transformed qualitative flight deck data were analysed and interpreted using Tableau software, lent context to the interview data, and were also compared to other existing AA SMS data streams to discover concordance and divergence.

Select Reflective Critiques
- LIT (Safety-II) observers are former LOSA (Safety-I) observers who, through diligent study and discussion, found that it took approximately six months to transform their knowledge, mindset, goals, and approach to data collection.
- *A priori* coding in the observation instrument carried the usual risks of such a coding scheme but was determined to be the most feasible given existing constraints.
- Iterative refinement of the observation instrument led to re-coding of the data, thus challenging validity.
- LIT team membership has waxed and waned since data collection began in early 2019, challenging continuity and cohesion. To remedy these concerns, LIT adopted a LOSA strategy of "cleaning" the data on a quarterly basis, excluding and re-coding as deemed appropriate.

3. Who would benefit from knowing about your work

This work is relevant to academic research / industry operational partnerships that wish to realize an organizational goal (in this case, development of a new safety data stream) via a rigorous methodological strategy. Mixed methods is an extremely versatile research approach that affords internal and external validity as well as numerous opportunities for secondary data analysis; additionally, mixed methods approaches can be designed to include frequent opportunities for establishing and maintaining concurrence between various datasets further validating the produced research conclusions.

4. The impacts or implications of the results or how they can be utilized

LIT methodology and methods serve as a case study for organizations seeking to incorporate a Safety-II/RE-informed data stream into their SMS. The designs of this project are specifically useful to those organizations which are unwilling or unable to completely abandon Safety-I work but are looking to improve operational safety through the incorporation of more modern approaches. Safety science researchers may also critique this work and build upon it in future research endeavors.

5. The wider scope or relevance of the work

Due to context-dependent features of the instrumentation, this work is immediately relevant to the aviation community, and indeed has already been considered for implementation by a large Australasian commercial airline. In a broader sense, however, mixed methods should be considered whenever resources permit such a robust research design.

Highlights of Recent Work

A Case Study in Instantiating RAG: American Airlines Learning and Improvement Team (LIT)

Christine JEFFERIES[a] and E. Asher BALKIN[a]
[a]*Department of Integrated Systems Engineering, The Ohio State University*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

Safety-II and Resilience Engineering professionals have many options for transforming industry safety programs, leveraging their expertise in the decision of which abstract theoretical model(s) to instantiate in a particular system. We present a case study and lessons learned from one large commercial airline's effort to utilize Hollnagel's Resilience Assessment Grid (RAG) in the transformation of their safety management system.

2. Summarize New results

American Airlines (AA), based in Fort Worth, Texas, USA, began to incorporate Safety-II and Resilience Engineering (RE) principles into their safety management system (SMS) in August 2018. The Learning and Improvement Team (LIT) chose to base their approach on the RAG model and traveled to Copenhagen to meet with Dr. Hollnagel in person. His advice during that visit dramatically shifted and since then has guided AA's Safety-II work: the team should not cut-and-paste the RAG model to structure their program but instead should take the principles of RAG and transform them to reflect AA's unique experience. With this bespoke approach, the LIT team sought and successfully achieved buy-in and funding from AA corporate leadership to pursue this avenue of study.

As such, Hollnagel's four abilities (Respond-Monitor-Anticipate-Learn) have evolved over the last two years to become LIT's four potentials (Learn-Coordinate-Adapt-Plan or "LPAC") as well as refining 20 proficiencies that describe instantiations of those four potentials. Furthermore, with ongoing support from the Cognitive Systems Engineering Lab (C/S/E/L) at The Ohio State University, LIT has developed a robust data collection methodology to discover examples of resilient performance and active graceful extensibility among pilots in the cockpit. While the process has been resource consuming and cognitively challenging, LIT has created a new, decision-driven data stream within their SMS and has facilitated fresh integrations within the organization between the safety department, training department, aviation human factors professionals, and others, as well as with outside of the organization with regulators and peer organizations.

In this presentation, we will share an overview of the evolution of the LIT program but will spend the majority of our time discussing the nuances of *how* the LIT team transformed Hollnagel's RAG model into LIT's LPAC model. We will share: our initial mindset as well as how it changed over the first two years; our deep dive into the epistemology of pilot work and how that was influential in developing LIT's model and methodology; finding balance between corporate goals, organizational capabilities, safety theory, and rigorous methodology; and critiques we've received since going public with the program.

3. Who would benefit from knowing about your work

While this work is directly relevant to the larger global commercial airline industry, any organization seeking to leverage RAG along their Safety-II journey would benefit from the lessons we learned along our journey. We demonstrate the importance of involving front-line practitioners in the process of instantiating an abstract theoretical model into practice, as a standalone effort and also in conjunction with existing systems.

4. The impacts or implications of the results or how they can be utilized

This is a novel approach to the instantiation of Safety-II/RE within an existing SMS as well as a unique, robust to data collection and analysis rooted in iterative, recursive, and abductive ethnographic methods. Analysis of the first 100 mixed methods flight deck observations and the first 8 qualitative pilot interviews has proved fruitful in restructuring of recurrent pilot training and captain leadership training, as well as sparking discussions with regulators. In the setting of AA's existing Safety-I-grounded SMS, the LIT program has inspired new considerations among safety and corporate leadership regarding resilience/optimality tradeoffs, how the theoretical framework underscoring which data and how data is collected can influence the collected data itself, and how learning from work as done can benefit the organization.

5. The wider scope or relevance of the work

AA published a white paper in June 2020 which described their Safety-II journey, including LIT goals, mindset, methodology, and preliminary results and insights. This paper has been shared widely among aviation professionals as well as other industries including aerospace, maritime, healthcare, and others. To organizations who are interested in instantiating RAG or another Safety-II/RE theoretical model, we offer our successful strategies as well as our lessons learned as we designed and implemented a Safety-II program.

Responses to overload as foundation for a resilient project and organizational health metric

Marisa R. G. BIGELOW, Lisa J. DOUGLAS, Alex MORISON, Michael W. SMITH, and Jordan M. HAGGIT
Mile Two, LLC, Dayton, OH, USA

ABSTRACT
It can be difficult to understand changes to team performance in the moment and at a larger organizational scale, especially when adapting to unique evolving pressures and opportunities. In some teams, situational awareness with respect to four responses to saturation - shedding load, reducing thoroughness, recruiting resources, and shifting work in time - may not be formally measured, or not measured collectively. To translate theory into practice and evaluate resilience at a project level, we are measuring the four responses to saturation on a regular cadence using a new project monitoring representation for communication and comparison at all organizational levels. Part of a larger project health metric, preliminary data for the four responses to saturation indicate an anecdotally significant increase in communication and awareness within the teams.

KEYWORDS
Situation awareness; engineering; saturation; metrics; project health; organization

INTRODUCTION

Most metrics used at the highest levels of an organization are often disconnected from the actual work being done and are narrowly defined to report on one aspect of a particular department or group within a much larger context. We are exploring a Project Health Measure that aggregates 6 metrics at the project level, enabling all the stakeholders at and above this level to have a shared understanding of the current state of the project as well as the ability to drill into each dimension that impacts the score. The Project Health Measure includes 1) team interactions, 2) artifacts produced by teams, 3) the quality of the software being developed, 4) how the project is progressing compared to the project plan, 5) a degree of customer satisfaction, and 6) a measure of project resilience. In particular for this context, project resilience uses the four dimensions defined by Woods and Hollnagel (2006) to capture the responses to saturation or overload: shedding load, reducing thoroughness, recruiting resources, and shifting work in time. With finite resources and evolving constraints, project teams and organizations are constantly navigating a complex trade-off space. A system's ability to respond to new demands can become saturated when workload bottlenecks exhaust the system's capacity to maneuver and adapt.

Organizations can greatly benefit from understanding how this construct can be measured, and resiliency increased, for the teams of people who do the work of the organization. Resilient teams are more effective and productive (Stoverink, Kirkman, Mistry & Rosen, 2020). However, the metrics used at most organizations are typically key performance indicators (KPIs) that reflect potential growth (e.g., qualified sales leads, employee engagement, NPS) but are by their nature abstracted and disconnected from the actual work being done. Project managements and software development teams also have KPIs to track things like bugs, budget, velocity, and volatility but these measures are typically focused on building the software and neglect some of the more intangible interactions that happen at the team level, like how people are getting along, is the customer happy with pace and progress, is the team producing what it says it will, or, importantly, how well the team is responding to changes thrown its way. It is also clear from searching the literature that organizational level metrics are usually employed when attempting to measure team level performance (e.g., Power & Waddell, 2004).

PILOT STUDY

We are currently piloting a subjective project health measure (referred to as Project Pulse) that aggregates six metrics at the project level from periodic survey measurements. The goal is to trace the critical decision points during the project in which the project lead and technical leads respond to increases in workload and significant changes in client direction. Project leads will evaluate the potential benefits and risks to document and support the decision of which class of response their team will follow.

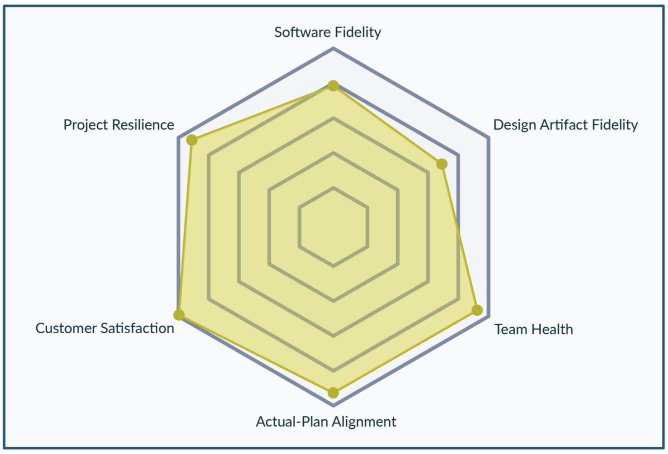


Figure 1. The radar graph for the Project Health Measure.

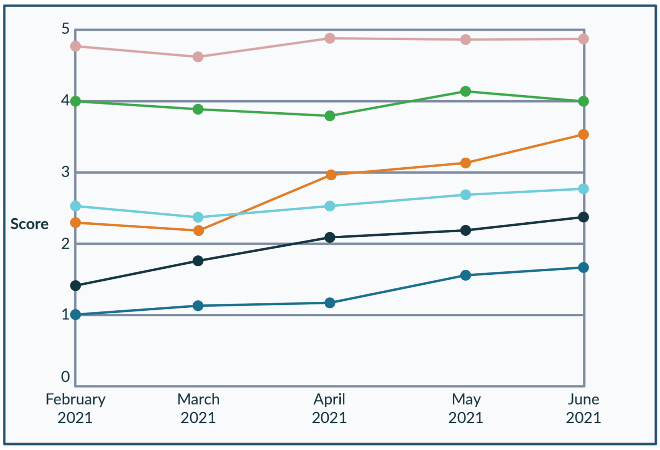


Figure 2. The time series for all six dimensions of the Project Health Measure.

Project Health Measure

Project Pulse measures six dimensions: 1) Team Health - team interactions, 2) Design Artifact Fidelity - research, design, and software artifacts produced by teams, 3) Software Fidelity - the quality of the software being developed, 4) Actual-Plan Alignment - how the project is progressing compared to the project plan, 5) Customer Satisfaction - customer or client contentment, and 6) Project Resilience - a measure of various responses to overload or saturation. The dimensions are represented as both a radar chart (Figure 1) for the latest measurement period and as a time series graph (Figure 2) for tracking changes over time. To support the project, we assembled a small team to build an online interactive application. However, the focus of this paper is narrowed to the measure of project resilience—four responses to saturation and overload described below. Specifically, the goal of the resilience metric within the Project Pulse pilot is to facilitate identification and understanding about the critical decision points during a project regarding the four responses. Would, or does, this approach to measurement via Project Pulse help a project lead better understand and communicate the current state of resilience? How does this awareness impact decision-making across the organization?

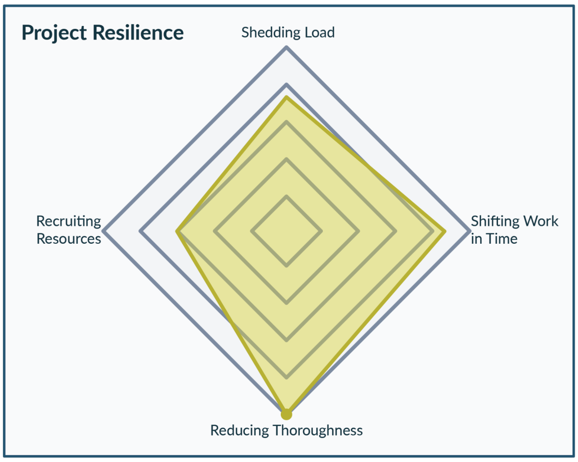


Figure 3. The radar graph for the Project Resilience measures.

Project Resilience

The four responses to overload (Figure 3) are coping mechanisms seen across various domains such as healthcare, aviation, etc. (Woods & Hollnagel, 2006). These categories were selected for the health metric as they can be tractably contextualized in the realm of software product development. For instance, shedding load refers to deprioritizing work, such as dropping tasks or requests. Reducing thoroughness reduces the amount of time allocated to complete the work, which could mean accelerating the pace of artifact generation relative to the team's underdeveloped work domain and problem understanding. Recruiting resources means to add people, capital, time, or material goods to alleviate workload pressure. Shifting work in time extends or delays the period in which work is completed, which could be exemplified as pushing tasks to the next sprint in an agile software development workflow. The four responses to saturation are further divided into tactical and strategic levels, where shedding load and reducing thoroughness are tactical responses useful in the short-term so long as priorities are well understood, and recruiting resources and shifting work in time are strategic with a longer-term focus. All four responses together give an awareness of the team's tendencies and explicit trade-off decisions.

Methodology

Our initial plan was to have project leads and technical leads report changes to workload and client direction by evaluating the potential benefits and risks to document and support the decision of which class of response their team will follow and subjectively reporting out biweekly. Table 1 shows an example of the resilience metrics questions from the biweekly survey. However, it was noticed for the first reporting period that the several project leads intended to include the entire collaboration team (research, design, and development) in the Project Pulse assessment. We are piloting with more consistent projects that we hypothesize would have similar tempos. However, if the pace and tempo dictate a different sensitivity, reporting also could be done ad-hoc. During and after the pilot period is completed, we will also query the project and technical leads about how they framed their decision-making for future, current, and past junctions. The aspect of risk is important to include when considering the consequences of actions taken by the project and technical leads. For instance, sacrifice decisions will inevitably arise with conflicting goals as in the case of multiple stakeholders or a shift in clients. Priorities must be balanced within limited resources, though it may not be explicitly apparent what opportunities the team may be giving up if only limited to considering tactical solutions. An affordance to articulate risk at the assessment time is provided by a free form text box in the Project Pulse application.

Table 1. The Project Resilience metric questions

Response to Saturation	Question
Recruiting resources	Did the team attempt to add resources of any type to the project (people, materials, capital)?
Shifting work in time	Did the team or any team member have to change any deadlines for any reason?
Shedding load	Did you or any team member need to reduce workload for any reason?
Reducing thoroughness	Did you, the team, or any team member reduce the thoroughness of your work (e.g., shortcuts, shortened deadlines) to improve efficiency?

We are currently evaluating two ways to use a Likert scale to measure the resilience responses. Using the example of *shed load,* one option is to define the scale from 1 to 5 where 1 means no load has been shed during the evaluation period (e.g., a week) and 5 means a lot of load has been shed. Alternatively, the *shed load* could be measured using a relative change where 1 means there is no change compared to the previous evaluation period and 5 means a lot of change has occurred. In an effort to understand which measure might be more sensitive in our environment, some project leads see their previous data entries when they fill out the questionnaire while others see a blank questionnaire every time data is submitted. Early data suggests it is more efficient for the project leads to see previous data so they can quickly evaluate whether something has changed since the previous reporting period.

Besides the individual dimensions' scores, project leads have two indicators they can switch that will communicate their current stance on project risk to upper leadership: action taken and support needed. Each question has both options to show if a particular aspect in one of the dimensions is at risk and if action has already been undertaken to address it. Initial data indicates a clear need for more communication around both options because the current context is unclear, particularly from question to question if the meaning remains consistent.

Additional perspectives from the team (e.g., quality assurance, technical leads, machine learning engineers, designers, software engineers) will weigh in on the decisions and provide contrasting views of the problem. For instance, task management tools for development and quality assurance such as JIRA, Gitlab, and Monday.com have specific metrics that act as a proxy for different types of responses. Changes in story point estimates and task deadlines can signal purposeful shifts in work further out in time by providing more resources and duration to execute the work. Adding contributors to a task indicates the team has recruited resources for that task. These actions in isolation, however, are not sufficient evidence for a particular type of tactical or strategic response. The action must be contextualized with the contributing factors and environmental conditions surrounding the decision through deeper feedback.

Unintended Consequences

Several unintended consequences were identified just before and at the pilot launch in the initial Knowledge Elicitation (KE) sessions with project leaders and other stakeholders. KE refers generally to interview techniques in the flavour of Flanagan, Klein, and others (e.g., 1954, 1989). First, we identified one question in the resilience dimension that could be problematic. If a responder wants to maintain the appearance that their work is of the highest calibre regardless of extenuating circumstances, the question, "Did you, the team, or any team member reduce the thoroughness of your work (e.g., shortcuts, shortened deadlines) to improve efficiency?" (Table 1. Question 4) might be at risk for a less than honest response. The culture at Mile Two–a highly collaborative systems and learning-forward culture–largely mitigates this risk but it should be noted that in many organizations this type of culture is not present and may increase the need for creating appearances. It is worth exploring how this question, as well as the other questions in the resilience dimension, may be related to the Efficiency-Thoroughness Trade-Off (ETTO principle) described by Erik Hollnagel (2009). His assertion that sacrificing thoroughness for efficiency is normal and we would benefit from understanding in what situations that is likely to happen in our own organization. Additionally, there could be privacy concerns both for this question and question 3 in Table 1. A responder may be hesitant to answer honestly because they may feel they are identifying themselves as the one who is underperforming, especially from an outside perspective. They were performing to the best of their abilities, of course, given the imposed pressures and constraints.

Second, across the questionnaire there is validity risk at the item (question) level within each dimension, including resilience. We have diverse project leads, projects, internal discussions, and how they interpret each question and make decisions about how to respond requires more research. We already know that some project leads find some of the scales (intended to be 1 = worst and 5 = best for all questions) ambiguous. One project lead mentions they will respond 3 because they think things are going well. Other project leads will respond with a 5 as they consider the project going as expected. One of the ultimate goals of Project Pulse is to understand performance across the organization and misalignment at the individual question level will negate this goal.

Third, we have found in early analysis that the word "resilience" does not resonate with most people in the same way it does with academics or the consumers of papers such as this. Here again, it is important to present the questionnaire in terms that are easily interpretable by the responder. In-depth research will help us create better questions and result both in better validity and reliability.

NEXT STEPS

One long-term goal is to include additional information or signals that can easily communicate risk - an indicator of whether the measured change over time is actually a risk. This relates to changes in the time series where the graph shows an upward or downward trend. For example, there are some times when shedding load is expected and not a risk to the overall health of the project; at other times this action may indicate a risk and should be communicated as such. The differentiation can enrich the communication and conversation about the project. Interview data are clear that consumers of the Project Pulse app need more communication around the *action taken* and *support needed* indicators. A longer-term goal will be to add functionality to include space for elaborating on the selected indicator, especially to help contextualize the responses relative to the risk identified.

The Project Health Measure can support anticipatory risk mitigation in the future by cataloguing various project paths and surfacing characteristics of the trade-off spaces. As mentioned earlier, the four responses to saturation can be separated into tactical and strategic categories. Although any of the four could be appropriate depending on the circumstances, the distinction between tactical and strategic highlights a fundamental trade-off between acute and chronic risks (Hoffman & Woods, 2011). Proactive strategies to manage these risks can help explicitly navigate the underlying Resilience-Optimality Trade-off that underpins the motivation behind the Project Resilience measure. Project management particularly struggles between more efficient, optimal planning and the need to be flexible, adaptive to changing dynamics. The various examples gathered through this effort can tangibly illustrate how diverse teams reasonably managed evolving trade-offs in retrospect.

IMPLICATIONS

The project metrics inform the next level higher in the organization at the portfolio level, in the pilot's case, which acts as a baseline for gauging the overall portfolio health. Traditionally, the portfolio directors relied heavily on the program managers and project leads' reflections on the various teams' progress. The health metrics would provide a similar point of comparison across the projects in collaboration with the leads' reflections. The evidence would be contextualized by each unique circumstance, but also generalized using the resilience responses to overload framework. The impact goes further to the executive level by then comparing across portfolios using the same basis at the micro-team level to abstract up to the macro-portfolio scale. The framework becomes a common language connecting the sharp and blunt ends of the organization, as well as demonstrating mindful team progress to other stakeholders (i.e., customers, operators).

The resilience metric applied to team project work would be of benefit to any organization that performs ongoing work using dedicated teams. The current test case focuses on a software development consultancy, though the work can be generalized to other domains. Organizations with multiple, distributed layers and teams could benefit from this work by applying the responses to saturation framework as a comparison metric. The metrics can enable teams and organizations to learn from previous project experiences to navigate future complications more effectively and create a larger learning culture. Locally rational decisions are contextualized in the flow of normal project work and thus, can be generalized using a universal framework for retrospective analysis. A corpus of project cases can inform future decision making by comparing similar circumstances and considerations. The different responses could also provide a prospective topology for alternative decision-making paths. Project leads, especially, can envision and simulate their responses to future pressures given the generic categories and relevant examples.

CONCLUSION

Work-level metrics can enable teams and organizations to learn from previous project experiences to navigate future complications more effectively and create a larger learning culture. Locally rational decisions are contextualized in the flow of normal project work and thus, can be generalized using a universal framework for retrospective analysis. For technology development organizations in particular, project and program management drive the product life cycles. Measuring resilience at the project level can provide important insight into team performance, identify issues and risk sooner, and highlight relationships between different layers of the organization and stakeholders. A corpus of project cases can inform future decision making by comparing similar circumstances and considerations. The framework and visualizations are in development now for contextualizing the trade-off space and comparing team performance around critical junction points.

ACKNOWLEDGEMENTS

We want to thank the Mile Two management for supporting us and for providing feedback in the development of this work. We also thank the participants of the pilot study as we continue to gather data and make improvements.

REFERENCES

Flanagan, J. C. (1954). The critical incident technique. *Psychological bulletin, 51*(4), 327.

Hoffman, R. R., & Woods, D. D. (2011). Simon's slice: Five fundamental tradeoffs that bound the performance of human work systems. In 10th International Conference on Naturalistic Decision Making, Orlando FL (pp. 5-31).

Hollnagel, E. (2009). *ETTO Principle: Efficiency-thoroughness Trade-off: Why Things That Go Right Sometimes Go Wrong.* Boca Raton, FL: CRC Press, Taylor & Francis Group.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on systems, man, and cybernetics, 19*(3), 462-472.

Power, J. & Waddell, D. (2004). The link between self-managed work teams and learning organizations using performance indicators. *The Learning Organization*, *11*(3), 244-259.

Stoverink, A. C., Kirkman, B. L., Mistry, S., & Rosen, B. (2020). Bouncing back together: Toward a theoretical model of work team resilience. *Academy of Management Review, 45*(2), 395-422.

Woods, D. D. & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton: CRC/Taylor & Francis.

Environmental, Health & Safety management system: Attributes and Barriers

Vanessa B. Bertoni[1], Ricardo Lecke[2], Ana Cristina Curia[2], Fernando Henrique Lermen[3], Vera Regina Pizza[1], Carlos Alberto Mendes Moraes[2]
[1]*Federal University of Rio Grande do Sul, Industrial Engineering Department, Brazil*
[2]*Unisinos University, Brazil*
[3]*State University of Paraná, Brazil*

1-Two sentence teaser that succinctly communicates why someone should participate and use your work

Environmental, Health & Safety (EH&S) integrated management systems are widespread in manufacturing industries, despite limitations regarding the system efficiency and performance, especially concerning technical and behavioural factors. The Functional Safety tool can collaborate with better results in implementing an integrated management system to prevent accidents and risks inherent to the productive processes concerning the protection of machinery and equipment.

2-Summarize New results

This study aims to identify attributes and barriers in the studies of integrating functional safety and EH&S with sustainability and management systems. Two main research questions guided this research: (i) Which attributes influence a sustainable implementation of an integrated management system? And (ii) What are the barriers to the performance of an EH&S management system?

A systematic literature review (SLR) was carried out in three phases: i) selection of studies on the topic, ii) bibliometric analysis, and iii) content analysis. This review, done through 702 searched documents from the Scopus, Science Direct, and Web of Science databases, resulting in 84 papers after exclusion criteria and 16 papers via the Snowball method. The methodological procedure structure is i) selection of studies on the subject, ii) bibliometric analysis, and iii) content analysis. This study uses the search string to apply the PRISMA protocol: ((Sustainability OR ("Environmental Health & Safety" OR "Environmental Health and Safety" OR EH&S)) AND ("Functional Safety" OR ("ISO 14001" OR "OHSAS 18001"))). The databases used were Scopus (310 documents), Science Direct (102 documents), and Web of Science (290 documents). After reading the titles and abstracts, 337 documents were excluded for not fit the scope. After the entire read of 106 documents, and the exclusion of 22 documents, 84 complete articles are part of the final sample of this research. In bibliometric analysis (ii), the journals, years of publication and the country of the first author were identified, Also, we built the relationships between the sample with EH&S, sustainability areas, application sector, and stakeholders. In content analysis (iii), we answer the RQ1, RQ2, and RQ3 based on relevant information identified during the sampled studies complete reading.

Finally, thirty attributes that influence the implementation of integrated management systems of EH&S and that use functional safety as a technological tool were selected. The barriers were distributed according to 6 dimensions related to the theme (Sustainable Supply Chain; Sustainability in the integrated management systems; Certifications and Standards to integrate management systems; Focus on people management; Environmental impact assessment issues; and Continuous industry improvement). The second place has six dimensions: the safety and integrated systems of environmental, health, and safety questions. In third place, functional safety covers three dimensions that most frequently address the discussion on accreditation, continuous improvement, and environmental impacts. Moreover, identifying common attributes that influence the implementation of EH&S management systems presents an inexistence of studies that evaluated the functional safety application as a tool.

Moreover, twenty-five barriers that use functional safety as a technological tool were selected, and distributed according to six dimensions related to the topic (Economic Issues; Human Issues; Industry ecosystem and its actors; Environmental issues; and, Standards and accreditors issues). The dimensions were grouped according to their sample incidence and themes related to the ISO 14001 items and OHSAS 18001 norms. Human Issues are inevitably aspects more necessary for the implementation of EH&S management systems, whether due to their training or awareness concerning the requirements of certification standards. Furthermore, this condition is no different in functional safety. Other dimensions that show barriers to implementing a management system are more related to external factors that are not on the company's management but can generate risks due to its direct influence on business. Finally, the barriers are directly connected to the expectations and needs of the interested parties and issues based on legal requirements and problems with relationships with authorities and supervisory/regulatory entities.

3- WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

The EH&S and functional safety integration support researchers, practitioners, and decision-makers in implementing management systems, promoting improvements in environmental, social, and economic aspects. The sustainability approach in these systems helps eliminate waste and increase profits; in particular, functional safety assists with reliability issues in the processes and products developed by manufacturing industries.

4-THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The risk management and environmental impacts prevention imply directly in the sustainability results, giving the industries a preventive institutional safety process. Besides, the optimized management maintenance and the search for what is valued by stakeholders can be guaranteed by the functional safety control in an integrated management system. When implementing integrated systems, manufacturing industries usually seek economic, environmental, and social resources optimization. Nonetheless, concerning the sustainability of an integrated management system, we consider that Functional Safety, coupled with EH&S, can impact different areas (i.e., aviation, maritime).

5-THE WIDER SCOPE OR RELEVANCE OF THE WORK

The main Functional Safety contributions in EH&S management systems are related to the concern with behavioral and technical aspects, aiming to build a responsible work environment. Also, these tools promote technical and quality training, promoting the development of standardized procedures and multidisciplinary collaborative workgroups, eliminating one of the previously identified barriers concerning the lack of training to implement training of employees in routines and management for leaders. Besides, these tools allow the implementation of technical and quality training, as there is a demand for the development of standardized procedures and multidisciplinary collaborative workgroups. These procedures can eliminate one of the barriers previously identified about the capacity to implement training for employees in routines and management for leaders.

Applying the Principles of Resilience Engineering to Small and Medium Sized Businesses

Simon GILL[a]
[a] *The Safeguarding Community*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- Small to medium sized business represent a significant contribution to the global economy and employment but minimal effort is made to build resilient performance.
- This presentation explores RE as applied to SMEs and summarises some of the resilience engineering initiatives which could be implemented with little investment by the SMEs themselves.

2. SUMMARIZE NEW RESULTS

Small to medium sized businesses (i.e. independent businesses with fewer than 250 employees) are a significant contributor to the global economy. In the UK for example, SMEs account for 99.9% of the business population, represent 61% of total employment and 52% of total business turnover[1]. Typically SMEs are smaller and more able to adapt to changes but are widely accepted to be more fragile than larger businesses. This is especially true in terms of a critical metric, financial security, as illustrated by a study at the start of the coronavirus pandemic when one third of SMEs globally expected to be out of cash without further support within three months and half of them within one month[2].

The coronavirus pandemic has exposed a need to build resilient performance of SMEs. SMEs were forced to adapt rapidly in multiple ways as the crisis evolved. Some of the challenges included a need to increase adoption of digital technologies, shifting the workforce to remote working, enabling flexible working for employees to care for family members, address fluctuating labour supply due to employees being unwell, permanent downsizing or furloughing of staff, minimising expenditure to maintain cash, seeking new customers and pivoting business models as previous markets contracted or disappeared and disrupted supply chains. Some of these effects were seen across all sectors, but some were more exposed to changing consumer habits or lockdown legislation, or gaps in governmental support packages. The impact of this is devastating. Despite governmental support measures, it has been anticipated that corporate insolvencies will increase by one- third worldwide by the end of 2021 compared to 2019[3].

There is a need to move beyond the traditional, minimal approach to risk management and explore how resilience engineering may enable SMEs to bounce forward from future shocks. SMEs also need to evolve in an environment of multiple stresses including the increased adoption of automation, artificial intelligence, and machine learning by larger competitors, evolving cyber security threats, fragile supply chains, environmental change and the resulting sustainability demands from customers and legislation. SMEs must also recognise that capitalising on opportunity can lead to shock and stress which must be identified, assessed and managed.

The majority of SMEs rely on simplistic risk management and business continuity planning approaches to identify, assess and treat threats and to prepare for specific risks. Many SMEs have also focussed on the elimination of slack and waste from their organisations through the application of popular business philosophies such as Lean and Agile, as well as in pursuit of maximum environmental sustainability. Such initiatives clearly have environmental, cost and efficiency benefits (and have important parallels with resilience engineering) but overzealous application can leave them dangerously exposed to the effect of disruptions. Building in an equivalent goal of resilient performance will require a different thought process.

The focus of this presentation is on practical solutions which can be picked up and applied by the owners and managers of SMEs, wherever possible without external support.

[1] UK Government Department for Business, Energy & Industrial Strategy (2020) Business population estimates for the UK and regions 2020
[2] Organisation for Economic Co-operation and Development (2020) Coronavirus (COVID-19): SME Policy Responses
[3] Coface (2020) Country And Sector Risks Barometer Q2 2020

The "four cornerstones of resilience engineering"[4] will be used to demonstrate theory and practice for SME owners:

Anticipating

There is still a place for risk management in the transition to resilient performance. However, risk management needs to change, reflecting operation in a so-called VUCA environment, where volatility, uncertainty, complexity and ambiguity are ubiquitous. We must educate those who own and run SMEs to move beyond a rudimentary, periodic scan for evident risks, assessed by likelihood and severity and placed on a risk register with mitigations. A modern approach is in constant review, actively seeks external information, embraces Strength of Knowledge, documents assumptions made and explores exposure to risk[5]. The process should recognise that surprises may still happen but support the development of preparedness for loss of critical functionality, irrespective of the risk source. More robust risk treatment including the use of targeted actions, experimentation and adaptive policies should be developed.

Learning

Learning from everyday work when things go well is a critical feature of resilient performance. In an average SME such activities have the potential to improve the efficiency of processes, optimising products for customers and improve the health and safety of employees and customers. SMEs should be exploring adaptations made by employees in the reality of their work compared to the work-as-imagined and work-as-prescribed. Active engagement of the employees in the process will improve motivation and a desire for the changes to be realised.

Embracing and encouraging a culture of sharing and learning from the coaching of owners and managers, to peer review processes such as action learning or appreciative inquiry to a robust, collaborative, open, peer-led investigation of where things have not gone to plan.

Responding

The organisation must be able to implement preparedness plans but also adapt in its response to unspecified disruptions. This requires consideration of longer-term structural changes, flattening hierarchies, cross-training staff and localising decision-making processes. Ensuring employee wellbeing initiatives are routinely used ensures that when facing disruption, individuals throughout the organisation have the capacity and capability to perform.

Monitoring

This requires first an understanding the critical functionality of the business and how it is delivered. Once this is developed, a suitable means of monitoring what internal and external factors influencing it should be developed.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Any owner or manager of an SME seeking practical, implementable actions to achieve resilient performance.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

SMEs adopting the principles of resilience engineering would be better placed to take advantage of opportunities and respond to disruptions in future. This presentation represents a practical means of achieving that.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Should SMEs widely adopt the principles outlines in the paper, the resulting resilient performance would make a significant improvement in how communities are able to bounce-forward from future shocks or stresses.

[4] Hollnagel, E. (2009). The four cornerstones of resilience engineering. In: Nemeth C., Hollnagel E. and Dekker S. (Eds.), Resilience Engineering Perspectives, vol. 2, Preparation and Restoration. Ashgate, Aldershot, UK

[5] Aven, T., (2014), Risk, surprises and black swans. Fundamental concepts in risk assessment and risk management. Abingdon/New York: Routledge.

The Impact of Learning from Everyday Work and Decision Making on Aviation Maintenance

Simon GILL[a]
[a] *The Safeguarding Community*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- Recent research on aviation maintenance decision making (including from Naturalistic Decision-Making) provides insights on how actions from resilience engineering may improve maintenance engineer performance and overall safety;
- This presentation summarises such action being taken in the offshore helicopter sector and provides recommendations for further work.

2. SUMMARIZE NEW RESULTS

Aviation maintenance engineers are routinely required to work around challenges in the maintenance environment to deliver aircraft safely and efficiently[1]. Such challenges can stem from organisational, environmental, individual, procedural and aircraft factors. The aircraft manufacturer Airbus sought to address the product factors by improving the design of the aircraft, tooling, procedures and documentation before the aircraft entered service in a process called Human Hazard Analysis (HHA)[2].

The process has since been adapted by the offshore helicopter community to explore the gaps between Work-As-Done on aircraft in-service with the Work-As-Imagined and Work-As-Prescribed during the design process. The process brings together designers from the Original Equipment Manufacturer (OEM) and maintenance engineers from the helicopter operators in workshop to focus attention on safety-critical maintenance tasks[3]. These tasks are analysed step-by-step to draw out errors, omissions, suggestions for improvement and potential error. The OEM is then able to take these ideas away, analyse and implement actions to address them.

Workshops have now been conducted on four helicopter types – representing a third of the offshore fleet – involving three OEMs and 11 Operators from four global regions. 818 safety-critical tasks identified by the OEMs were analysed and, in addition to many ideas for improvement, 256 potential errors were identified of which 89% of the errors were deemed low risk with effective control measures and low consequence, 9% were moderate and 2% were rated high. Actions have been taken by all the OEMs, with improvements to design, tooling and documentation have been implemented or are in progress. Bell, Sikorsky, Leonardo and Airbus have collaborated to create a "Standardised Approach" to this process which should be ready to share by Q3 2021. Further results from this proactive safety analysis will be shared in the presentation.

These results reinforce the underlying premise of resilience engineering that we have much to learn from everyday work but the HHA process, although very effective, is highly resource intensive in its current form and can only be applied on safety-critical systems. Alongside this consideration by the OEM, operator focus on human factors in the hanger aims to optimise the conditions under which maintenance is carried out. However, inadequacies in the maintenance environment require engineers to make quick, safety-critical decisions under uncertainty. Supporting maintenance engineer in their decision making is therefore a critical activity. Recent studies of aviation maintenance engineers as experts, guided by Naturalistic Decision Making (NDM) provide some insights into improvements that can be made.

It is generally accepted that the maintenance manuals are as a starting point in undertaking a maintenance task. An important consideration is when maintenance engineers decide to deviate from those manuals. Such deviation occurs due to the identification of a "*better way, more effective or efficiently, and a safer way*" of

[1] Tsagkas, V., Nathanael, D., & Marmaras, N., (2014) A pragmatic mapping of factors behind deviating acts in aircraft maintenance, Reliability Engineering and System Safety, vol. 130, pp. 106-114
[2] Lawrence, P. and Gill, S. (2007), Human Hazard Analysis, Disaster Prevention and Management, Vol. 16 No. 5, pp. 718-739
[3] Gill, S. (2019) Use Of Human Hazard Analysis To Enhance Resilient Performance Of Helicopter Maintenance Systems, 8th REA symposium Embracing resilience: Scaling up and speeding up, Kalmar, Sweden

conducting the task. Such deviations are a threat to aviation safety but they also contain opportunities to identify problems in design, tooling, documentation and improvements in procedures for the benefit of all[4].

In the same study it was found that 45% of maintenance engineers routinely found technical errors in maintenance publications. 78% of participants also demonstrated some level of agreement in understanding how the process to amend publications worked. This indicates that either lots of amendments are in process or they are simply not being reported. "In either case, inaccurate or ineffective publications provide a pathway to deviant behavior".[5]

The role of affect in maintenance decision making was also addressed. Anger and contempt created action tendencies for risk taking, whilst pride was found to have a protective effect[6]. The decision-making model employed was concluded to be rational for trouble-shooting and intuitive in numerous examples regarding personal safety (e.g. use of a harness, stepping on the top rung of a ladder, working in a fuel tank without PPE) and in a decision to cease a maintenance operation in bad weather[7].

These studies illustrated other factors influencing decision-making including situational awareness, the age profile of maintenance engineers, time pressure and the time taken to process changes to documentation.

This presentation will illustrate the results of these studies which should be translated into action by aviation organisations:

- We must look to alternative feedback mechanisms built into the fabric of the maintenance system, facilitating the efficient and effective sharing of ideas and innovations by maintenance personnel through their own organisations and to the OEM;
- OEMs should facilitate a quicker, easier process to report documentation errors and omissions, and provide more feedback on action taken;
- We should seek an opportunity to bring unapproved maintenance aide-memoires (or so-called black books) into the open to allow the opportunities therein to be realised and risks eliminated. This should be done openly and encouraged by MROs as part of their Safety Management Systems.
- The innovation of maintenance engineers and positive deviance from manuals should be bought into the open, encouraged and facilitated;
- Consideration should be given to the application of Aeronautical Decision-Making (ADM) for maintenance and the implementation of Competency Based Training;
- Maintenance engineer training should build on many of these recommendations including hazard awareness, the shared understanding of safety in the organisation and innovation and positive deviance from manuals. Building in an explanation of decision-making but also a space in the training to allow practice of real-life situations to positively reinforce safe judgement.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

All maintenance professionals would find value in this work. It addresses theory but also translation of theory into practice in offshore helicopter maintenance, but the principles and recommendations can apply to wider industries.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

This work can be readily applied to inform practice of maintenance professionals.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

This work seeks to improve the safety of offshore helicopter maintenance but application beyond can have an influence on safety in other fields.

---

[4] Bannister-Tyrrell, A.R. (2020), Uncertain outcomes: Traversing the breach between innovation and violation in aviation maintenance, Doctor of Philosophy Thesis, University of Newcastle, Australia

[5] Ibid

[6] Naweed, A. & Kingshott K. (2019) Flying Off the Handle: Affective Influences on
Decision Making and Action Tendencies in Real-World Aircraft Maintenance Engineering Scenarios Journal of Cognitive Engineering and Decision Making, Volume 13, Number 2, June 2019, pp. 81–101

[7] Hemingway, D. (2019), Aviation Maintenance Technician Decision-Making, Doctor of Philosophy Thesis, Walden University

Highlights of Recent Work

Use of Smart MMO Virtual Mega Worlds to Accelerate Innovative business Models and Expand Complex decision making in Organizations

Lia A. DiBello Ph.D, David Max Lehmann Ph.D. PE, MBA CRM[b], E.S. Chamberlain III and Daniel Rohrlick B.S.
Workplace Technologies Research Inc. and Applied Cognitive Sciences Labs Inc.

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

For decades, our team has explored accelerating expertise through a specific kind of "gamification", which emphasizes unstructured tasks with high stakes "non-negotiable" goals and opportunities for rapid trial and error iteration. This results in the development of deep domain expertise.
Recent work incorporating the use of Massive Multi-player virtual worlds with advanced behaviour tracking as a way of increasing the power of this method and scaling it has yielded some unexpected benefits.

2. SUMMARIZE NEW RESULTS

BRIEF OVERVIEW OF METHODS

For decades, our team has explored accelerating expertise through a specific kind of "gamification", which emphasizes unstructured tasks with high stakes "non-negotiable" goals and opportunities for rapid trial and error iteration. This results in the development of deep domain expertise. This approach – although new – has deep roots in long standing research on expertise (e.g., Chi, Glaser and Farr, 1988) and the cognitive mechanisms involved, which we have discussed in other work (DiBello 2011; DiBello 2019; Hoffman et. Al 2014). In 2007, we incorporated the use of Massive Multi-player virtual worlds with advanced behaviour tracking and instant feedback as a way of increasing the power of this method and scaling it. The difference between these applications and our previous attempts were size and complexity. The number of people in the environment and the number of moving parts that could be in play (such as technologies and impacts) were much greater.

In this paper, a specific case study will be explored which is illustrative of the potential. The topic here is the unanticipated outcomes; these same worlds can be used to iteratively rehearse company strategy options that are too risky to rehearse or simulate other ways. Specifically, the *immersive participation of actual workers, customers or stakeholders executing proposed strategies or even company designs "inworld" allows all to see if the proposed plan will survive reality – and more to the point – contact with human actors.* There are a number of implications for NDM research.

The case itself involves a Mining portfolio company. Normally a mining company goes after "mega deposits", which can take up to 20 years to discover and qualify and which can cost over $2-3 Billion to develop before any gain is realized. The eventual yield from the deposit, which can last 40 years or more, pays for the upfront cost. On the other hand, the company must continually offset the cost of mining itself with fluctuating mineral prices and cost of fuel, labor and other factors.

As these large deposits are far and few between, companies are looking for ways to make money from smaller deposits and speed up the discovery and start up process. This project involves using our technology and an A.I. discovery system that eliminates much of the drilling to reduce the cost of discovery to a few Million and reduce the time to profitable operation to one year. The project involves 11 smaller mines and deposits that would not be possible to develop in a cost-effective way using older methods. In a virtual world platform, they will be managed as one mega deposit, managed with one business strategy, even though not co-located in real life. Operational options are being "rehearsed" in digital twins of the actual mines while workforces for each mine are trained together on technologies shared across mines in a multi-player, cloud resident virtual world. The financial and operational "control room" for whole portfolio is virtual and shared by all the properties.

NEW FINDINGS

1. Not only are individuals' capabilities expanded, but the team as a whole experiences an "extension" in their shared ability to expand their problem solving and "what if" thinking by manipulating large elements of the business in the "mega" world.
2. The platform is going beyond more passive "digital twins" where large factors can be manipulated and observed, but where participates cannot "go in" and affect the simulated environment by their own individual behaviour, assumptions about deployment, or decisions, such as would be the case in real life.

3. Micro and Macro feedback, delivered together helps participants "connect the dots" between individual behaviours and their own impact on larger events, increasing their capability to mentally simulate eventualities.
4. A new business model is emerging. Early on, it became clear the business being modelled could not have been possible without this kind of technology because the cost of failure was too high. The technology removed this risk and opened the door to this opportunity for a new model of doing business.

PROJECT PROGRESS

As this abstract is being written, the platform and it's major elements are assembled and the "what if's" of the business are being rehearsed. Because of the reduced costs, riskier elements with higher pay-off potential are being introduced at this stage, such as exploration and design of higher value mines (normally costing close to $1Billion), a plan to go public, and an aggressive marketing play to potential investor/buyers.

NEW COLLABORATIONS

The project is bringing together a widely inter-disciplinary team and introducing NDM methods of research to these teams. These include: Gold investors, mining companies in far reaching places, refineries, other NDM researchers, Geophysicists and business people interested in new portfolio management models not profitable without advanced expertise and the technology to support them. It is also drawing the attention of those interested in the Future of work and workforce development.

NEW STUDIES

A potential study that could emerge as this project continues is the ways that technology expands the cognitive capability of individuals and teams by allowing them to externalize problem solving in an immersive reality-analogous virtual environment of a problem space that they can control and manipulate. The role of immediate feedback, problem space exploration, and compressed time can be examined to explore the various impacts on cognitive capability in a number of decision making domains.

WORK TO TRANSFER KNOWLEDGE INTO PRACTICE IN DIFFERENT SETTINGS

This project is one of many where we are deploying a "mega world" to manage a complex set of problems to achieve complex goals. There are many other settings under consideration, such as "circular economies" in Africa, responsible forestry and mining of the Amazon, leadership models for colleges, and new agricultural models.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Anyone who wishes to know the impact of advanced technology on developing the capability for extended decision making, collaboration and examine the role of "rehearsal" for understanding complex problems.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The data collected and the methodological journey in this study can inform other studies of accelerated learning, especially as they relate to research on teams solving complex problems.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

We see relevance to: workforce development, the role of technology at every level of the organization, the development of new business models and the effective deployment of A.I., onboarding and training.

References:

Chi, M. T. H., Glaser, R. and Farr, M. L. (Eds.) (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum.

DiBello, L., & Missildine, W. (2011). The future of immersive instructional design for the global knowledge economy: A case study of an IBM project management training in Virtual Worlds. *International Journal of Web-based Learning and Training Technologies*, *6*(3), 14-34.

DiBello, L. (2019). *Expertise in Business; evolving with a changing world.* Ward, P., Schraagen, J.-M., Gore, J., and Roth, E. *The Oxford Handbook of Expertise: Research & Application*. Oxford: Oxford University Press.

Hoffman, R.R., Ward, P., Feltovich, P.J., DiBello, L., Fiore, S.M. and Andrews, D. (2014). Accelerating Learning and Expertise: Concepts and Applications. Boca Raton, FL: Taylor and Francis.

Resilient Analysis Grid: a quantitative approach of healthcare provider’s perspective during COVID-19 pandemic

Vanessa B. Bertoni[1], Natália Ransolin[2], Priscila Wachs[3] and Angela Righi[4]
[1]*Federal University of Rio Grande do Sul, Industrial Engineering Department, Brazil*
[2]*Federal University of Rio Grande do Sul, Civil Engineering Department, Brazil*
[3]*Federal Institute of Education, Science and Technology of Rio Grande do Sul, Brazil*
[4]*Federal University of Santa Maria, Brazil*

1- TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

The COVID-19 pandemic unfolding requires an investigation of healthcare providers' resilient performance at the front line of patient care during services' disruptions imposed by the 'new normal'. A Resilience Analysis Grid (RAG) virtual survey could be an alternative approach for unveiling healthcare systems' resilience potentials to manage pandemic disturbances while preserving patient safety and occupational health and safety.

2- SUMMARIZE NEW RESULTS

This exploratory study has the following research question: how healthcare providers perceive their hospitals' resilience performance? Adopting a resilience engineering perspective helped investigate the relationship between the four resilience potentials (respond, monitor, learn, and anticipate), patient safety, occupational health and safety from the provider’s point-of-view.

Due to the pandemic constraints to physically collect data and the broad research scope, a quantitative approach was needed to capture and analyze data related to healthcare providers' assessment of their institutions. In practice, a questionnaire based on the resilience potentials was developed to evaluate healthcare institutions' quality and safety of work conditions during the current pandemic. The survey was virtually applied to a snowball convenience sample, in which invitations were sent through email and social media to potential participants.

The data collection occurred from June to August 2020. The ethical precepts were respected during the study development. The inclusion criteria were healthcare providers working in Brazilian hospitals during the COVID-19 pandemic. From the 111 valid responses, 8% of the participants work in a private hospital, 62% work in a public hospital, and 30% work in both. 40% work in the Intensive Care Unit (ICU), 13% in the patient ward, 9% in the Emergency Department and 39% in other hospital units. And, 47% of the participants are nurses, 14% physical/respiratory therapists, 10% doctors, 5% others.

The instrument was divided into two topics, the first one related to the participants and their hospital characterization (e.g. profession, private or public hospital); the second one, based on a 5-point Likert scale, associated with the four resilience potentials. Each resilience potential was analysed in relation to nine major categories: (i) patient safety; (ii) previous experience; (iii) training; (iv) procedures (standard operational procedure, protocols, flows); (v) communication; (vi) resources (people, equipment and material); (vii) built environment (workspace); (viii) occupational health and safety; and (ix) resilience potential efficiency.

For instance, considering the training category, the proposed questions were: (i) responding - “Professionals are receiving training concerning the demand of infected patient”; (ii) monitoring - "Professionals are receiving training to monitor the status of the institution facing the pandemic and use it to adjust their response"; (iii) anticipating - "Professionals are being trained to anticipate their response toward different situations"; (iv) learning - "Lessons learned from past events are being applied in training to cope with COVID-19".

This study corroborates the principle of interconnected performance between the four potentials to support the system's resilience. The survey results suggest that when providers perceive their workplace as safe, they also consider patients safe. The primary outcome is that learning and responding potentials received more emphasis than the monitor and anticipate potentials. For example, the question "lessons learned from positive experiences from your team or other institutions contribute for coping with the pandemic" presented the highest mean (4,43). This point can be explained by the fact healthcare providers are constantly aware of new COVID-19 evidence while coping with demand instability. Indeed, during the pandemic, institutions do not have enough time available to promote mechanisms of anticipation.

3- WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Our findings might be of interest to researchers and practitioners of organizational resilience aiming to understand and collaboratively incorporate healthcare providers' feedback into the organizational operational procedures. Understanding which aspects in the work-as-done (WAD) by healthcare providers affect patient safety and occupational health and safety can support managers in decision-making. A RAG virtual survey is an alternative tool to support resilience management when collecting data in loco is not possible, and the study has a broad range (nationwide).

4- THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

In addition to the difficulties commonly experienced by the healthcare sector, such as the hospitals' overcrowding and the scarcity of fundamental resources, situations such as the high demand for personal protective equipment and its availability contribute to services' disruptions and chaos in healthcare systems. This study presents a step forward in promoting organizational resilience performance by practical and theoretical insights, such as evaluating the four resilience potentials and their impact on patient safety and occupational health and safety during a pandemic. Moreover, a practical tool that can be used to analyze how resilience develops over time is presented.

The pandemic's learning curve has been especially steep since most workplace problems, and possible solutions can be reapplied to healthcare systems worldwide. The dissemination of effective practices for building a more resilient WAD in healthcare systems would be the ultimate benefit of this research.

5- THE BROADER SCOPE OR RELEVANCE OF THE WORK

Brazilian healthcare systems are close to collapse because of the rise in COVID-19 cases, as demand for coronavirus-infected-patient care increases. For instance, ICUs occupancy rates in Brazil are over 100% (as of March 2021). Warnings concerning healthcare providers' health and safety are gaining attention since the unsafe and stressful conditions they are working under are getting common. Although this study points out that there is a need for a deeper investigation of the resilience potentials to monitor and anticipate healthcare institutions' performance during the pandemic, findings deserve future research, including interviews and a mixed-method approach. Nevertheless, the tool RAG survey application has effectively achieved the study aim since it was possible to be adapted for a virtual environment.

Highlights of Recent Work

Developing a learning organisation

Pedro FERREIRA [a], Nippin ANAND [a], James KWASNY [b] and Nick PETERSON [b]
[a] *Novellus Solutions*
[b] *Allied Pilots*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- Learning requires effective communication that fosters organisational cohesiveness, as opposed to contributing to its fragmentation.
- Effectiveness of communication requires moving from one-way channels towards multi-way, bottom-up and top-to-bottom, as well as peer-to-peer conversations.

2. SUMMARIZE NEW RESULTS

Project Confidus received initial funding from the Scottish Oil & Gas Institute of Technology (OGTC) and emerges from the growing perception that most organisations invest substantial resources in reporting-based communication systems and yet, draw little benefit from such investments. It focuses on shifting from essentially one-way and highly fragmented communication practices, towards multi-way conversations. These conversations provide a much wider context on issues and therefore, generate enhanced potential for learning. The engagement in conversations produces reflection and sharing of knowledge with and between the various organisational levels. The use of dedicated technology aims to overcome time and geographical barriers and foster improved hierarchical, as well as peer-to-peer exchanges across remote assets.

Initial findings have shown that this approach leads to a significant lowering of the threshold for detection, as issues can be discussed and explained over the conversations before they become effective operational and/or safety problems.

The similarities of Confidus with the ongoing work by the American Airlines Learning and Improvement Team (LIT) has inspired a cooperation, which beyond exchanging on the different solutions and approaches taken, aims at exploring further usage of the data being generated. Qualitative analysis and Artificial Intelligence solutions are being considered to expand the potential of the diversity of perspectives and of situated expertise that can be gathered through conversations (in the case of Confidus) and pilot debriefs (in the case of AA LIT) towards improved learning across the organisation.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Researchers can benefit from the methodological approach developed, but mostly industry members can benefit from the achievements of the applied solutions that are being implemented. Aside from the participation of members from the AA LIT, the plan is to involve representatives from organisations that have been involved since the early research stages of Confidus and others that joined in the course of more recent commercial oriented work.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

In addition to a more effective application of situated expertise into problem solving, the ultimate goal is to uncover the potential for generating significantly improved perception of shifting trends across the organisation and support business decision-making and strategies.

5. The wider scope or relevance of the work

Improved maturity and accumulated knowledge can significantly improve proactiveness and adaptive capacities across the organisation.

Highlights of Recent Work

Team Decision Making in Designing Software for Joint Cognitive Systems

John FLACH, Alex MORRISON, Marisa BIGELOW, Jonny BUTLER and Ariel SWIFT
Mile Two LLC, Dayton, OH, USA

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- Although Mile Two considers Cognitive Systems Engineering (CSE) to be essential for modelling/analyzing cognitive work, we realize that cooperative teaming with multiple other disciplines is necessary for designing and developing software products that will support that work.
- The challenge is how to turn a collection of diverse disciplines into an effective design team, where all are empowered to contribute alternative perspectives and to participate in design decision making.

2. SUMMARIZE NEW RESULTS

CSE provides an incredibly useful framework for understanding a work domain and the cognitive challenges that impact the ultimate quality of performance in a joint cognitive system. However, those trained in CSE rarely have formal training in design and those trained in design disciplines rarely have a background in CSE. Thus, creating a common ground for effective teamwork can be challenging. For example, the value of initial time spent doing thorough work analyses is not always appreciated by designers and developers who are anxious to start building wireframes, prototypes, and code. Additionally, designers and developers may have difficulty interpreting and translating the products of work analysis into more traditional design artifacts. Without establishing common ground, the danger is that teams will iterate around the design cycle (Figure 1) without converging on a satisfying solution. This is particularly dangerous in waterfall or bucket brigade style collaborations where there is significant danger that the insights generated by the CSE analyses early in the design process have all 'sloshed out of the buckets' by the time a product is produced at the end of the line.

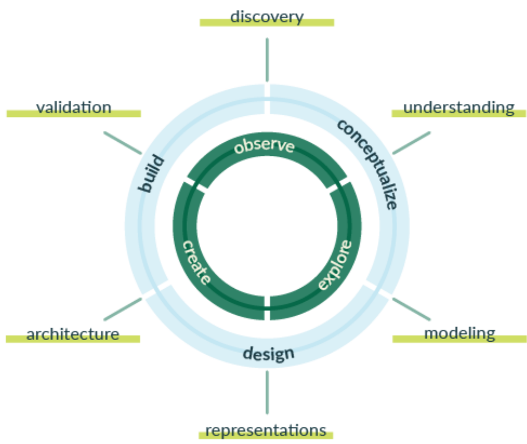


Figure 1: Design Cycle.

Mile Two has begun experimenting with alternatives to more traditional design approaches that seems to be meeting with some success. Here are three important aspects of our approach: 1) Recognizing the value and limitations of alternative perspectives; 2) Cross training; 3) Co-creation.

Value and Limitations

Academic debates between various disciplines (e.g., HF vs. CSE) tend to be framed in either/or terms to determine which perspective is correct. We believe a first step toward effective collaboration is to reframe the discussion in terms of both/and, where the value and limitations of alternative perspectives is explicitly recognized. Figure 2 illustrates how several important disciplines are complementary in terms of spanning the dimensions that ultimately determine the quality of a design.

Cross Training

Mile Two has invested a significant effort to cross train our designers and developers in CSE. This has involved a CSE bootcamp that has been taught by David Woods, and weekly CSE working group sessions to give people hands-on practice in doing work analysis and generating various CSE artifacts as depicted in Figure 1. We are also encouraging our people who have been trained in CSE to experiment with more traditional design tools (e.g., Adobe XD for generating wireframes and prototypes) with guidance from our people who have been trained in UX/UI disciplines.

Co-creation

The most significant change is to the design processes. We have shifted from a process that tended to be serial to a process that is more parallel. In the serial process, much of the early knowledge elicitation and work analysis was done by CSE people, who in turn worked with UX/UI designers to create wireframes, that were then handed off to developers to do the coding. In contrast, in the parallel process, all the disciplines are involved in all stages of the design. Thus, UX/UI designers and developers join the CSE people for many of the knowledge elicitation sessions and the team works together to model the work in terms of functional decompositions, state transition diagrams, scenarios, etc. Then the entire team explores alternative interface representations through sketching and ultimately the generation of wireframes and prototypes. The coding is done largely by developers, but there is close collaboration with the other team members to verify that the ultimate product is consistent with the design specifications. At various stages of this process stakeholders (e.g., potential users or domain experts) are engaged, and hypotheses and assumptions are revisited and revised to reflect a new or deeper understanding of the work to be supported. The unique feature of our approach is that diverse disciplines contribute at every stage of the development process.

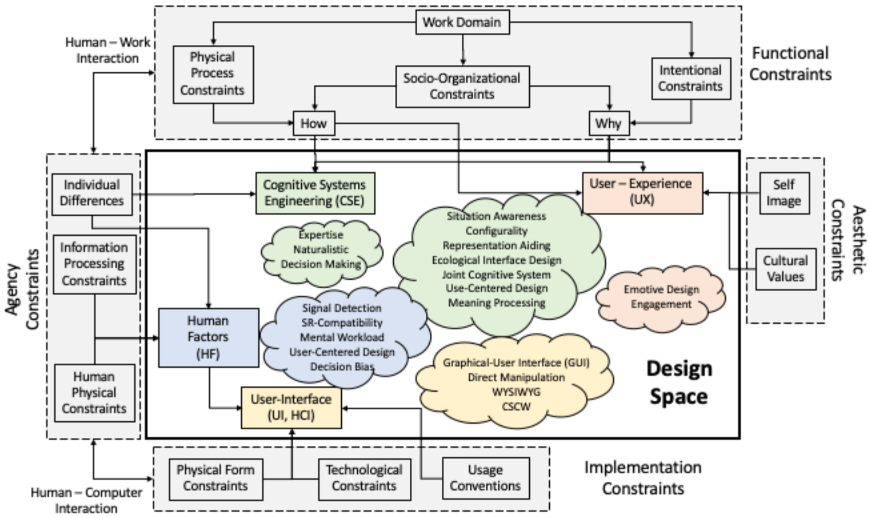


Figure 2. Multidisciplinary perspectives within the design space.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

This work should be of particular interest to people who are interested in applying a CSE perspective to the design of software to support cognitive work. The focus is on an integration of CSE principles into a practical design process that converges on innovative solutions to cognitive challenges.

This also has implications for academic programs. Although it is impractical to design programs that cover the full span of the design skill set, programs can be designed to heighten awareness and appreciation for the values of multiple perspectives in a collaborative co-creation process.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The ultimate impact of this work is to increase the likelihood that the principles of cognitive science as inculcated in a CSE framework will inform the design and development of software that supports cognitive work. In discussing our efforts toward developing more effective design processes, we will be candid about both our successes and difficulties. Perhaps, others can build on our successes and avoid some of the pitfalls that we have experienced.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

In addition to improving design processes, we believe the new processes tighten the coupling between design and cognitive science in ways that may feedback to increase our insights into the nature of joint cognitive systems and decision making under natural, operational conditions.

Learning from the positive – suggestions for potential improvements to maritime safety management approaches

Johannah Olsson and Gesa Praetorius[b,c]
[a]Stena Recycling AB
[b]Linnaeus University, Kalmar, Sweden
[c]University of South-Eastern Norway, Borre, Norway

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

This article presents findings from a recent study on Safety Management Systems (SMS) in the maritime domain. We present how safety perceptions between shore-based and ship-based differ, and suggest a set of Resilience Analysis Grid-based questions that can be utilized to complement current SMS approach, thus a contribution to closing the gap between work-as-imagined - or prescribed - and work-as-done.

2. Summarize New results

The study presented in this article aimed to investigate how regulatory demands have shape safety management and safety management systems (SMS) in shipping companies. A case study focused on five companies operating in the Swedish maritime industry was carried out. Ten semi-structured interviews with respondents in management positions on shore and onboard were carried out during spring 2020. Originally the case study was supposed to include observations as well, but these were suspended due to COVID-19. The interviews were transcribed and analyzed with the help of the four resilience cornerstones; respond, anticipate, monitor, and learn.
The results show that the companies, as well as the industry itself should be considered as complex systems, in which an additional layer of complexity is added by international and national rules and regulations. The work put into managing safety has become of such complex nature that for the most part, safety work is put into administrative tasks to ensure compliance with regulations. The results also highlight concerns among the management onboard that more time is dedicated to formulate routines and procedures than to actually making the work safer for those performing it. To compensate for the experienced gaps between work-as-imagined, or prescribed, and work-as-done questions based on the Resilience Assessment Grid (RAG) are presented. These can serve as a complement to the current safety management strategies, and may also give raise to a deeper cooperation between shore-based and ship-based personnel in shipping companies. This may enable a participatory approach in which safety management can be shaped bottom-up with as onboard personnel contributing essential knowledge on needs and preconditions to ensure safe operations while also taking the complexity of the system into account.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

This study highlights the complexity of implementing safety management in a high-risk domain. While the findings are based on a study in the maritime domain, they may be interesting for safety professionals in other high-risk domains.

The findings of this study can benefit those aiming to increase safety of work within complex socio-technical systems. Knowledge of what the current characteristics of the safety management and SMS are within these companies makes it possible to identify risks with current approaches, as well as to identify potential improvements towards a more proactive SMS approach.

THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

It is suggested due to the result found in this study that a Safety II approach to safety management, along with the use of resilience engineering to develop and enhance the domain's adaptability, can serve as a complement to the current safety management approach. Being able to adapt, respond and manage various unforeseen situations is a way of ensuring safety in operations even in complex socio-technical systems. The resilience assessment grid, RAG, is suggested as a tool to be developed to be usable in the participating companies. The RAG could serve as a tool to taper the gap between WAI and WAD, as well as to provide input to the development of indicators other than accidents for improving safety. Furthermore, it could also facilitate learning from everyday operations and what is going well – the everyday successes in everyday execution of tasks involved in operations.

4. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Factors such as complexity, vastness of current SMS and the want for equipment, procedures and routines that better fits the purpose, as well as the gap between WAI and WAD, form incentives for a complementary approach to safety management within the studied companies. This calls for a further study within other companies within the domain to be able to state whether this is the case within the domain as a whole or not. If so, Safety II offers a new dimension to the current safety management by recognising and enhancing the adaptability of the humans in the systems. Resilience engineering offers the possibility to manage and develop the potentials for the adaptability that is needed to handle the variability in complex socio-technical systems.

This then also have implications for other complex socio-technical systems that also have safety management systems that are becoming complex and procedures and routines that are deemed unfitted to the purpose.

Highlights of Recent Work

Understanding the applicability of resilience practices: proposed approach in the investigation of solutions to support societal resilience

Matthieu BRANLAT[a], Asbjørn Lein AALBERG[a] and Martina RAGOSTA[a]
[a]*SINTEF Digital, Trondheim, Norway*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- To improve its capacities, an organisation might want to learn from tools and practices successful elsewhere.
- Project ENGAGE proposes a path to understand the conditions for applying such practices in another context.

2. SUMMARIZE NEW RESULTS

International recommendations provided in the United Nations Sustainability Goals and the SENDAI framework emphasize the value of public participation in all phases of disaster management activities (from preparedness to learning from events) in order to build more resilient societies. But how do organizations leading these activities, such as authorities and emergency organizations, go about this? How do they identify potential approaches or even specific tools to better involve populations under their responsibility? Can they, for instance, look at innovative and successful approaches in other areas and directly apply them in their location? Can they, also, recognize and better support local initiatives that emerge from the population's own efforts? These are the questions project ENGAGE aims to answer over the course of its 3-year span.

The primary objective of ENGAGE is to find *solutions* to involving civilian populations in disaster management and describe the conditions necessary for their successful implementation. We use the term *solutions* in a broad sense: it encompasses tools, methods, practices, procedures, and strategies. The approach that we adopt in ENGAGE is based on understanding the specificities of a given location. We consider that the most effective solutions will necessarily be related to how societal resilience manifests in a given location. We ask, for example, what makes certain solutions viable in one local context but not (or only partially so) in another? To make the link between the solutions and societal resilience, we first consider what elements of the local context are factors that support or hinder societal resilience.

Here a distinction is helpful: on one hand, there are aspects of society on which one can act upon directly (e.g., income level) – these are **target** aspects; on the other hand, there are aspects that are part of the local context and cannot easily be changed (e.g., culture) – these are **contextual** aspects. This distinction between target and contextual dimensions allows us to state the underlying idea of the project: a solution aims to have an impact on a population, but its impact is influenced (positively or negatively) by the context of use.

This idea can be illustrated using an example of solution from the current COVID-19 situation. Most countries around the world have implemented some form of regular briefings about the situation. The ways in which briefings are carried out varies (e.g., regularity, content), but the main objectives are usually to keep the population informed about rates of infection, recommendations, and vaccination progress. In other words, such a solution aims to increase the population's risk awareness. We investigate in the project how such awareness in a population is indeed related to societal resilience. But the success and impact of such communication vary, though, in part due to contextual factors such as trust in authorities, an aspect that is typically associated to societal resilience. Such trust is usually described as cultural (varying across countries, cultures and communities) but can also be related to a particular moment in a location's history (e.g., specific government in place). In any case, while risk awareness in the population is indeed something authorities might try to improve directly (i.e., target), increasing trust in authorities might be a longer-term objective and is difficult, if not impossible, to address directly.

Interestingly, the distinction between target and contextual aspects is not always so sharp. Over time, improvements in target aspects might translate into improvement in contextual aspects; and solutions, while more targeted in certain regards, might indirectly aim to impact contextual aspects. For instance, using our previous example, establishing better communication practices between authorities and the population in times of crisis might, in the end, improve trust in authorities.

These are the kinds of factors of societal resilience, solutions for involving the public and complex dynamics, that ENGAGE aims to investigate. The project aims to improve knowledge by identifying and relating such elements and dynamics, and to improve practices by providing enriched descriptions of solutions that authorities and emergency organizations can implement to better support societal resilience in local contexts. In the end, a solution involving a population in disaster management might be successful in one location but only be implementable elsewhere under some conditions. The overall ideas discussed here, which underly the work in ENGAGE, provide a path to understand why such a solution might be successful and how sensitive it is to its context of use.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Resilience researchers and practitioners can find inspiration in the approach presented in the paper, in particular: researchers interested in the multi-faceted nature of resilience; people interested in "best practices" and their limitations.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The work presented challenges the notion and frequent use of "best practices" by emphasizing: (1) the situated nature of resilience, and (2) the impact of context on the successful application of practices or tools.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The project and paper proposed focus on societal resilience and disaster management. However, the approach described, i.e. the detailed characterization of solutions and distinction between contextual and target aspects, are relevant to other domains. Indeed, the underlying questions around implementability of solutions and multi-dimensional nature of resilience apply also to organizational practices in critical infrastructures, for instance.

For the context of resilience engineering, and the conceptual and pragmatic investigations of resilience, the distinction between target and contextual aspects should be informative. It might help in clarifying the discussions around topics such as best (resilience) practices or safety culture.

PLEASE LIMIT YOUR SUBMISSION TO TWO PAGES OR LESS

Highlights of Recent Work

Rethinking mass casualty distribution – Embedding a resilient hospital selection algorithm into a mass casualty distribution simulation model

Sheuwen Chuang[a], Chia-Hsin Cheng[a]
[a]Graduate Institute of Data Science, Taipei Medical University, Taipei, Taiwan

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

Decision on the distribution of mass casualties to appropriate hospitals is a critical practice in mass casualty incident management. Distributing victims by considering stakeholders' resilience potentials rather than just applying the short-distance and medical-care adequacy policy may need to be planned. The study developed a hospital selection algorithm for patient distribution and a mass casualty distribution simulation system embedding this algorithm to analyse these kinds of decisions.

2. SUMMARIZE NEW RESULTS

- Brief overview of methods

First, the algorithm development used nation-wide hospitals' capacity data (205 hospitals), mass casualties' data of the Formosa Fun Coast Dust Explosion (FFCDE) event (499 victims), the OpenStreetMap data from the disaster scene to individual hospitals, and the received mass casualties data over time of the 36 initial receiving hospitals, as well as the interview data responding to the influx of burn patients from some initial receiving hospitals.

Next, modeling the mass casualty distribution system used the discrete event simulation technique in the Anylogic software. The model was designed into seven modules beginning from the initiation module of a mass casualty incident, parameters setting, rescue and triage, waiting for ambulances, hospital selection, evacuation, and output module. Finally, the algorithm embedding into the simulation model was verified by cross check between two model designers and validated by using the victims' distribution data of the FFCDE scenario.

- New findings

A basic hospital selection algorithm is developed, which is an equation constructed by three elements: transformed value from a route distance between disaster scene and each hospital, a digital number selected from the medical-care adequacy matrix based on patient's triage level, and it is adjusted by the number of received patients over time in each hospital, and mobilization capability determined by the ratio of hospital's staffing and amount of critical beds and adjusted by average ED overcrowding rate of last two years. A hospital score is the sum of the standardized value of the three elements. Every victim has a list of scores for all possible sending hospitals according to their triage level. The highest score indicates the best decision to transport the victim to a proper healthcare facility.

The algorithm with and without the mobilization ability was tested by the simulation in the FFCDE scenario to analyze whether the mobilization ability affects casualties' distribution. The study found no significant difference in total transportation time between the two kinds of algorithms, but the algorithm with mobilization ability brought more severe victims to the severe-level hospitals and more moderate victims to the medium-level hospitals than without.

Further, comparing with the distribution results of the 499 victims in theFFCDE, the average total transportation time of running ten times of the simulation model using the algorithm with the mobilization ability under the FFCDE scenario is longer than the actual event. The number of receiving hospitals in simulation has eight more hospitals than in the actual event. The study found that one different severe-level hospital, eight different medium-level hospitals, and seven different general-level hospitals were between the simulation and reality. Besides, severe-level hospitals in simulation received 87 fewer victims than in actual event, moderate-level hospitals in simulation received 16 more victims than in actual event, general-level hospitals in simulation received 37 more victims than in actual event.

Moreover, the simulation can find an optimal combination of all resources, such as number of rescuers, number of emergency medical technicians, and number of ambulances for each scenario through comparing and assessing parameters setting.

- Project progress

  The study developed a simulation system to support the analysis of resilient mass causality distribution.

  The simulation can provide suggestions for mass casualty distribution in different scenarios based on a short distance and medical-care adequacy policy, as well as the possuble mobilization ability.

  This model can run one disaster occurring in multiple scenes. It provides a friendly UI, animation of transporting victims to hospitals, and varied analysis figures and tables. The mobilization ability is still in doubt. The study will refine the current algorithm and advance the modeg to be capable of running more than one disaster at a time, and considering the effect of self-transport minor injured victims to hospitals.

- New collaborations

  The study currently collaborates with the Situation Center of Ministry of Health and Welfare to obtain more domain knowledge and data, and practical experiences.

- New studies

  An App of mass casualty distribution adopting the hospital selection algorithm and using real-time hospital data is developing to support the exercises of mass casualty incidents.

- Work to transfer knowledge into practice in different settings

  The combination of using the simulation model and the App can be transferred to the local governments and the regional emergency medical operation centers (REMOCs) in well preparedness planning for local and regional resource allocation, as well as to the local Fire departments and hospitals in responding to mass casualties of varied scenarios both in exercises/drills and a real event.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

The local rescue team and hospitals would benefit from knowing the proposed optimal distribution results derived from the simulation model to prepare their response plan in varied simulations of mass casualty incidents. Moreover, using the App, the casualties distribution information in real-time at a real event can be shared among rescue teams, hospitals, and REMOCs, to proactively respond to the influx of mass casualties into hospitals over time.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

Overall, the information generated by the simulation model and the upcoming App could support disaster planning, training, exercises/drills design, and responding to MCIs for the emergency medical services system as a whole to improve disaster readiness.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The wider scope of the series of studies is to build up system resilience of the emergency medical services. The information generated by virtual and reality through simulation models and the mass casualty distribution App supports human decision-making in routine training and disaster practices during MCIs.

Resilient healthcare during the COVID-19 pandemic: a case study in a Brazilian hospital

Natália RANSOLIN[a], Carlos Emilio Stigler MARCZYK[a], Rafael Parmeggiani GERING[a,] Tarcísio Abreu SAURIN[b,] Carlos Torres FORMOSO[a]
[a]*Universidade Federal do Rio Grande do Sul, Civil Engineering Department, Brazil*
[b]*Universidade Federal do Rio Grande do Sul, Industrial Engineering Department, Brazil*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

The COVID-19 pandemic has challenged the resilience of healthcare services in several countries. This exploratory study presents the lessons learned by a leading private hospital in Southern Brazil, using the lens of resilience engineering.

2. SUMMARIZE NEW RESULTS

Eight semi-structured interviews were conducted, involving caregivers (nurses and physicians) of front-line patient care and managerial positions and an infrastructure manager. Interviews were conducted in December 2020, a period in-between pandemic waves of COVID infections in Brazil. Additionally, researchers visited the hospital and went through the main patient flows (COVID and non-COVID), in order to observe the changes in the built environment to cope with the pandemic. Data from the interviews and observations provided a basis for the development of a FRAM model that represented the couplings between the major resilient practices to address the pandemic and the functions affected by them. The study was approved by the hospital's ethics committee.

Six lessons for supporting resilient healthcare were devised from the case study. These lessons were associated with the four resilience abilities (anticipate, monitor, respond, and learn). The lessons were grouped according to their implications to the design and the operation phase of health services and facilities. The former is framed as Work-as-Imagined (WAI) for health services and Built environment-as-Imagined (BEAI) for healthcare facilities. In turn, the operation phase is framed as WAD and Built environment-as-Done (BEAD).

The lessons learned for the healthcare systems´ design phase correspond to the resilience ability of anticipation. Major threats such as pandemics are expected, and therefore hospitals must be at least partly prepared for them. The main lessons for the WAI/BEAI are related to promoting flexible workspaces to accommodate functions and their variability along the system´s lifecycle. In turn, the lessons for the operation phase of healthcare systems are mostly related to the resilience abilities of monitoring and responding. Monitoring might point out gaps between the BEAI and the BEAD, which might trigger responses to threats unanticipated in the design phase. The lessons for the WAD/BEAD are mainly related to developing internal capabilities for the best use of available resources (e.g., physical attributes and technological devices), which includes their quick reallocation when necessary. Similarly, maintaining a multi-skilled workforce to cope with demand surges is highly recommended. The need for collaborative work to support quick responses to unexpected events was also highlighted in the case study.

The match between capacity and demand has been a key challenge during the pandemic, and the main lessons learned refer to this point. Reports of collapsed health services during the pandemic indicate that the main bottleneck to match capacity to demand is the lack of qualified staff rather than lack of physical infrastructure, even though this latter has also been severely challenged.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

This report aims to contribute to health workers at care and administrative positions in collaboratively finding strategies to cope with unexpected events. Designers of the built environment and health services can benefit from our findings. Furthermore, this work can be useful for the investigation of the relationship between resilient healthcare and evidence-based design.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

This case study investigated how the COVID-19 has affected healthcare resilience, highlighting implications for services and the built environment. Since the lessons learned are not limited to a specific hospital unit, being of interest to several healthcare services. As the pandemic is still unfolding, the lessons learned are subject to revision. Nevertheless, the case study of a leading private hospital offered insight into the challenges posed by the pandemic. Despite the substantial financial slack and a well-structured organization, the hospital of the case study has been severely challenged by the pandemic, reinforcing the need for long-term thinking in healthcare systems' design and operations.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Several research works have been conducted attempting to understand the COVID-19 pandemic causes and consequences. Further studies will focus on updating the lessons learned as the pandemic evolves and retrospective in-depth studies after it subsides. In this respect, a more structured data collection approach will be used, mapping the patient flows as well as the main flows of the supporting resources (e.g. drugs) throughout all hospital areas in order to identify the resilient performance in each process stage.

How to foster participatory innovation during crisis management?

Samira BOURGEOIS-BOUGRINE[a], Nathalie BONNARDEL[b], Sylvie TORDJMAN[c], Silke SCHAUDER[d], Todd LUBART[a]

[a] *Université de Paris and Univ Gustave Eiffel, LaPEA*

[b]Aix-Marseille Univ., PSYCLE, InCIAM

[c] Rennes 1 University, Centre Hospitalier Guillaume Regnier. INCC- Université de Paris and CNRS UMR 8002

[d] Picardie Jules Verne University - CRP-CPO

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- This action-research provides a proof-of-concept that participatory innovation during crisis management cycles can be used to facilitate the improvement of organizations' resilience
- Lessons learned from the roll out of participative innovation during COVID-19 pandemic (formalized as an AFNOR-Spec)

2. SUMMARIZE NEW RESULTS

- Brief overview of the background and the aim

  Back in 1976, M. F. Weiner wrote an article in the journal *Medical Economics* entitled "Don't Waste a Crisis — Your Patient's or Your Own." Weiner meant by this that a medical crisis can be used to improve aspects of personality, mental health, or lifestyle. Thus, our research project aims to promote the emergence of new solutions to workplace-related problems raised by COVID-19 and establish a foundation for on-going processes to deal with such disruptions and crises in the future. Unlike the "*top-down*" approach to innovation management, "*bottom-up*" management is implemented to directly solicit employees and/or users, whose experiences and perceptions in the field will serve as foundation to the emergence of new ideas. This action-research was carried out as part of project SURVIE (SURpassing the Virus with Innovations in Emergence), funded by the National Research Agency in France.
- Method

  A participatory innovation action has been deployed in five domains (transport, maternities, start-up business, university, specialized hospital). Concretely, this consisted of the implementation of an idea management system, intended to encourage the emergence of staff and/or users' ideas and to facilitate their exploitation during the COVID-19 crisis. The approach involved 3 kinds of partners:
  - Academic team (authors) with diverse background: ergonomics, psychology, psychiatry, creativity and innovation.
  - Experts in management from AFNOR group: International Institute for Competency Development – HR21 –(ICD-HR21).
  - 5 organizations: a specialized hospital in Britanny (Centre Hospitalier Guillaume Régnier), a network of Maternities in eastern France (CoPéGE - Champagne-Ardennes- Lorraine region), a university faculty (Aix Marseille University), SNCF-Transilien (RER Paris region transport), OCUS (a French start-up in the international cultural industry sector).
- Results
  - The implementation of a transversal approach, analyses performed on data gathered in the 5 organizations and return on operating experience (through interviews, questionnaires) allowed the definition of specifications, which were formalized into the form of an AFNOR Specs.
  - The AFNOR Specs includes a set of organisational principles and guidance on the characteristic of an effective innovation climate as well as 5-step framework for the roll out of a participative innovation action during crisis management.
  - Future uses of these specifications in diverse crisis and domains will contribute to the improvement of the participatory innovation AFNOR Specs.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Public and private organisations whatever their size and the domain

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The participatory innovation AFNOR Spec provides insights on how to launch, foster and channel ideas generation, facilitate concepts' implementation and promote the learning process. This optimized process results in new and useful ideas with economic value and improved social well-being.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The participatory innovation framework could be a part of an organisational crisis continuity plan.

Highlights of Recent Work

A tale of two "cities": comparing the State of Ohio and The Ohio State University's ability to enact massive change

Michael F. RAYO[a], Samuel S. MALLOY[b], and Elisabeth D. ROOT[c]
[a]*Department of Integrated Systems Engineering, The Ohio State University*
[b]*John Glenn College of Public Policy, The Ohio State University*
[c]*College of Public Health, Department of Geography, The Ohio State University*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

We will share our front-row-seat perspective of how the State of Ohio (population 11.7 million) and the community of The Ohio State University (population of ~100,000) generated new capabilities to meet the demands of the pandemic. It reveals the barriers and facilitators for change (i.e., adaptation) amid a climate in which the need for change was urgent and unambiguous.

2. Summarize New results

Although we are all researchers, and this work has a research component, it was not and is not primarily research work. We were contracted by the state of Ohio, and later by The Ohio State University, to invent, implement, and participate in the Comprehensive Monitoring Teams for both organizations. The CMT's were designed to gather data from many disparate sources, periodically brief decision-makers, and influence COVID-19 response decisions and actions. Neither of these organizations had Incident Command (IC) or Intelligence, Surveillance and Reconnaissance (ISR) functions that they could leverage or extend for their CMT. Both had to be created from scratch *during a global pandemic*. With our team having expertise in public health (Root), public policy and IC (Malloy), and resilience engineering and ISR (Rayo), we advised first the state, and then the university on how they should envision, resource, and support these fledgeling teams. Our decisions were designed to inculcate resilience into these reconfigured and reconfiguring organizations by designing for resilience in the new CMT's. This included being flexible but focused on what capabilities should be recruited into these *monitoring* functions, what interconnections needed to be constructed to the various *response* functions, and the creation of analytic tools that would facilitate *anticipation* of the future trajectories of the virus. Although we did not set up any formal structures for it, we also tried to be the *learning* function for the teams, quickly understanding what had and had not been working for the World Health Organization and Centers for Disease Control, other countries, other states, other universities, and within each CMT.

Among the obstacles and surprises (both situational and fundamental) that we encountered included:

- Most of our help was not wanted - we recruited a large number of volunteer organizations that were willing to donate hundreds if not thousands of person-hours of work, and were experts in the functions that were required for these new functions. Very little of this volunteer work was accepted. None of it was accepted after the first two months.
- Since CMT membership was never permanent, and all resources were on loan, there was continuous pressure from everyone's previous commitments to minimize effort on the CMT. This remained true even as the pre-pandemic structures and processes proved inadequate to mitigating the spread of the virus. For the state, this resulted in dissolving

the team as the Autumn 2020 surge began. At the university, dissolving the team has been averted by strong leadership and real connections to university's COVID-19 response.

- Any intervention that directly interfaced with the public was met with stiff resistance. Conducting primary research and analysis (e.g., speaking directly with Ohio citizens, OSU students) was either discouraged or disallowed.
- It took very little to stymie progress. To do anything new or different required that multiple people work outside of their role to coordinate groups that had historically never coordinated. Many proposed efforts failed to gain sufficient traction.

We learned the following about how these Societal-scale Systems challenge assumptions that we make about our Industrial-scale Systems.

- There is no one really in charge. The Governor is not really in charge of the public, and definitely isn't in charge of the legislature. There is no expectation of adherence or compliance to rules or regulations relative to what our expectations are of industrial-scale systems. Even at the university, the president really doesn't control student, staff or faculty behaviors but there is more control.
- There is HUGE diversity of perspective, with very little (if anything) that is universally understood. There might be some shared common ground within one camp or another, but there is very little between them.
- There is HUGE variability in information access and knowledge. This is about facts on the ground, or the underlying mechanisms (e.g., science) related to the phenomenon of interest. Variability across the university community was less than that of the state, but still was non-trivial and meaningfully large.

- **Who would benefit from knowing about your work**

Anyone serious about supporting or designing resilient systems will benefit from knowing about this work. This is especially true of those working with increasingly large systems, with the ultimate being societal-scale systems (SSS). The phenomena that we observed are present in other systems - the size, design, and heterogeneity of the system merely amplified it.

- **The impacts or implications of the results or how they can be utilized**

If RE/NDM is serious about wading into this new scale of system, we need new ideas, new collaborators, and quick dissemination of our findings about the vastly different dynamics and pitfalls of societal-scale systems. We will once again be outsiders, but there will once again be fellow travelers who have been making these societal-scale systems go. Or, at least, are committed and at least periodically successful in making them go in a specific direction!

- **The wider scope or relevance of the work**

It is clear to us that the dynamics that we encountered and worked through are prescient facsimiles of what will be faced for all future societal issues. These include global climate change, widespread poverty, substance abuse, structural racism and many others. Creating sustained adaptability in government and societies will require rethinking of the interactions between people and their governments. The future will echo the lessons that this pandemic is teaching us.

Highlights of Recent Work
Implications for World Model Development Based on Perceptual Cues and Strategies Used by an Experienced Pilot to Negotiate Wires Obstacles

Katie ERNST[a], Laura MILITELLO[a], and Robert COPELAND[b]
[a]*Applied Decision Science, LLC*
[b]*U.S. Army Combat Capabilities Development Command Aviation and Missile Center*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

- Helicopters and other rotorcraft fly almost continuously in close proximity to hazards and are just beginning to incorporate advanced automation that will rely on dynamic world models.
- Our work highlights the perceptual cues and strategies of human rotorcraft pilots in navigating complex airspace, and identifies opportunities to support human knowledge and resilience to improve overall mission effectiveness and safety.

2. Summarize New results

Helicopter operations in degraded visual environments such as rotorwash-induced brownout, fog, clouds, and battlefield obscurants are significant source of risk. The rotary wing aviation community has dedicated over two decades of research to addressing this challenge due to the high number of accidents, injuries and deaths in degraded visual environments (Crowley et al., 2000; Vreeken, 2013) . A significant advance is the inclusion of a world data model, capable of representing the complexity of the helicopter operating environment, for use in navigating safely through free space (Whalley et al, 2014). A voxel-based model ingests point-cloud data from radar and/or LIDAR-based sensors and produces a three-dimensional model of empty, filled, or unknown space that can be used for navigation through the field of obstacles toward a distant location (Whalley et al, 2014). This in turn offers opportunities to reduce the workload associated with rotary wing navigation and pilotage, as is evidenced by the Army's Mission Adaptive Autonomy program (Whalley et al., 2016). Rotary wing operations are generally high workload due to the complexity of helicopter flight dynamics and close proximity to vertical obstructions. Human pilots develop perceptual cues and strategies for navigating through this complex airspace, allowing them to operate efficiently and resiliently as mission and flight changes occur.

Our present work is designed to elicit and represent the perceptual cues and strategies that helicopter pilots use to navigate wire obstacles. Wire obstacles include electricity and telephone wires and their support towers or poles. Helicopters often fly at low altitudes where wires present a significant hazard. Risk of striking a wire and damaging or disabling the aircraft is exacerbated because wires are nearly invisible to the naked eye while in flight. In this presentation, we will revisit existing work on three-dimensional world models representing empty, filled, and unknown space and their relevance to the wire obstacle challenge. We will then recount the story of a helicopter pilot who survived a wire strike. Using this story, we will articulate important cues and strategies that this pilot used to initially negotiate the wire obstacle and then ultimately recover, fly, and land the damaged helicopter. This story exemplifies the unique sources of resilience afforded by experienced helicopter pilots as they strive to execute the mission even in marginal conditions. The perceptual cues and strategies identified using this story have implications for important aspects of helicopter pilotage and navigation to be incorporated into dynamic world models.

3. Who would benefit from knowing about your work

This work has implications for the research and development community in human-machine teaming as the community develops support that supplants traditional human sensing and sensemaking with technology-mediated sensing and sensemaking. This work is also of value to those researching and developing automation and autonomous functions in all modes of transportation.

4. The impacts or implications of the results or how they can be utilized

These results can be used directly in the design of dynamic world models for use in rotary wing aircraft systems design. A broader audience for these results includes those developing advanced air mobility products and systems that intend to leverage the airspace in and around urban and suburban areas as alternate means of transportation. Unmanned aircraft supervisors directing aircraft in these complex environments will effectively be operating in degraded visual environments given the requirement for technology to mediate their ability to sense and avoid adjacent obstacles and air traffic. System design that incorporates the cues and strategies of experienced pilots to navigate these complex environments will leverage and extend that knowledge in new and meaningful ways.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

This work has relevance to the design of automation and autonomous functions in all modes of transportation. The navigational challenge of rotary wing aviation in close proximity to obstacles in degraded visual environments is directly relevant to those developing systems for autonomous ground vehicles. In addition, this work would be useful for those developing advanced air mobility concepts intending to leverage the airspace in and around urban and suburban areas as alternate means of transportation. Unmanned aircraft supervisors directing aircraft in these complex environments will effectively be operating in degraded visual environments given the requirement for technology to mediate their ability to sense and avoid adjacent obstacles and air traffic. World models that sufficiently support the joint performance of humans and automation will lead to safer, more resilient and effective system functioning.

REFERENCES

Crowley, J. S., Haworth, L. A., Szoboszlay, Z. P., & Lee, A. G. (2000, June). Helicopter pilot estimation of self-altitude in a degraded visual environment. In *Helmet-and Head-Mounted Displays V* (Vol. 4021, pp. 304-315). International Society for Optics and Photonics.

Whalley, M. S., Takahashi, M. D., Fletcher, J. W., Moralez III, E., Ott, L. C. R., Olmstead, L. M. G., ... & Conrad, B. (2014). Autonomous Black Hawk in Flight: Obstacle Field Navigation and Landing-site Selection on the RASCAL JUH-60A. *Journal of Field Robotics*, *31*(4), 591-616.

Whalley, M. S., Takahashi, M. D., Mansur, M. H., Ott, C., Minor, J., Morford, Z., ... & Schulein, G. J. (2016). Flight test results for a new mission-adaptive autonomy system on the rascal juh-60a black hawk. In *American Helicopter Society 72th Annual Forum. West Palm Beach (FL)*.

Vreeken, J. (2013). *Helicopter flight in a degraded visual environment*. (NLR-TP-2013-559). European Aviation Safety Agency. https://www.easa.europa.eu/sites/default/files/dfu/Final%20Report%20 EASA.2011.02.pdf

ACKNOWLEDGEMENTS

Work by K. Ernst and L. Militello is funded by the U.S. Army under contract NNA13AB88C Subcontract Number AS19-01059.

Modelling driver decisions to improve port traffic management during critical situations

Ross Owen PHILLIPS[a], Samaneh REZVANI[b], and Ben RUTTEN[c]
[a]*Norwegian Center for Transport Research (TØI), Norway;* [b]*DEMO Consultants, Delft, the Netherlands;* [c]*Ben Rutten Technology Research, the Netherlands*

We explore the idea that treating human decisions according to meaning processing approaches to cognition can lead to more satisfying management of real-world situations. We do this by attempting use the model of Flach & Voorhorst (2020) to improve the results of system theoretic process analysis (STPA) of traffic management in a busy port area of the Netherlands.

The study is ongoing and will interest researchers and practitioners looking to account for human factors in systems engineering, port and traffic managers implementing technological solutions, and those interested in ecological psychology and STAMP paradigms. The study is part of the EU project SAFEWAY.

The real-world challenge

Trucks delivering containers to a shipping terminal often need to pass by entrances to other terminals in order to reach their destination. If one of the bypassed terminals must close due to IT failures, high winds or capacity issues, trucks waiting to deliver containers can build up outside its entrance and block the passage of other trucks. To solve the problem, port management use road signs to divert incoming trucks headed for closed terminals to a holding area situated away from the entrance of the closed terminal. In practice, many of the truck drivers ignore the diversions and still head for the closed terminal. The existing road layout is such that trucks cannot be physically channeled to the holding area. Management want to know how to manage the situation by other means.

This real-world challenge concerns how to control driver actions in a complex sociotechnical system (Vicente 1999). It is becoming increasingly popular to frame such challenges using the Systems Theoretic Accident Model and Processes (STAMP; Leveson, 2012) paradigm, and analyse them using system-theoretic process analysis (STPA; Leveson & Thomas, 2018). In STPA, the analyst begins by defining **losses** to be avoided in the system (e.g. delay to container deliveries), system **hazards** that can could result in losses (e.g. roads blocked by trucks), and **system constraints** that must be controlled to avoid hazards (e.g. trucks must not block roads). The analyst then maps purposeful actions and information flows between all actors and technologies in the system, before identifying for each actor **unsafe control actions** that if performed could violate system constraints. Finally, **loss scenarios** are identified to explain why people or technologies could act "unsafely". These scenarios can then inform system improvements.

The methodological challenge

A central question in STPA is: What would make human or technological actors in the system perform **unsafe control actions**? Human actors are central to any sociotechnical system, but the STPA handbook provides little guidance on how analysts should model human perception and action (Leveson & Thomas, 2018). Several attempts have been made to elaborate this aspect of STPA (Thornberry, 2012; Montes, 2016; France, 2017), of which the extension described by France

(2017) has been popular. France's (2017) model of human control as summarized by Figure 1.

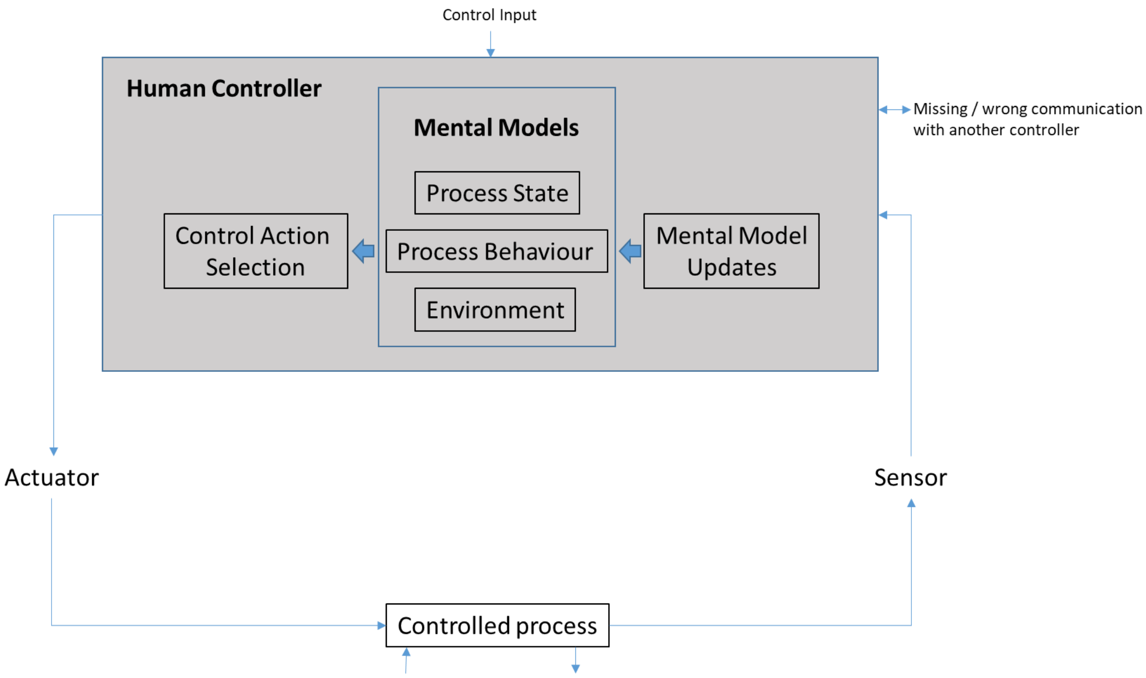


Figure 1. Model of human control used in "Engineering for humans". Adapted from France (2017).

In line with this model, France (2017) includes a set of questions analysts should ask when considering scenarios that could explain unsafe control actions. These are summarized in Table 1.

Table 1. Prompts analysts can use to consider scenarios that could cause unsafe control actions and system losses (France, 2017)

Main question	Prompts
How did the operator choose which action to perform?	• How did operator make their decision? • What goals were there? • Multiple/conflicting goals? • Designer vs. operator goals? • Available action alternatives? • Was the decision making skill-based (was a familiar action attempted when not appropriate?), rule-based (which rules were applied?), or knowledge based (what mental model of system or process?) • How were mental simulations of potential actions and consequent decisions influenced by time pressure, fatigue, workload etc.
What does the operator know or believe about the system?	• What mental models («cognitive representations of the world») did the operator have about: - process state? Model mismatch, mode error, mode change triggers - process behaviour? "Why is it doing that?", "what will it do next?", "what actions are available in current mode?" "How will inputs affect system behaviour?" - environment? Familiar or novel? Weather, complexity of situation, beliefs about other controller actions and mindsets
How did the operator come to have their current knowledge or beliefs?	• What difficulties are there in creating and maintaining mental models? • What properties of system or sensor information are most salient? • What feedback and input does the operator expect or think they should monitor? • What do they not expect or not monitor? • Effort for operator to access needed information?

The advantages of France's (2017) extension to STPA is that it models human control in a way that is both aligned with STPA terminology and accessible to engineers, such that they might account for human factors in their designs. Given the growing popularity of STPA, however, the modelling of human control has received little attention in the academic literature. In particular, there have been no attempts to conduct STPA while modelling human control using concepts that are explicitly aligned with models of human perception and control action in real-world environments, such as naturalistic decision-making (Klein, 2017), ecological psychology (Gibson, 2015) or dynamic human control (Jagacinski & Flach, 2009). These approaches have recently been collected, developed and espoused in a single model – the **meaning processing approach** to human cognition – but no attempts have been made to extend STPA and account for human control using this approach (Flach & Voorhorst, 2020).

The meaning processing approach

According to the meaning process approach, cognition is not located within the controller, but emerges from the ongoing interaction between a person and their ecology. In addition, cognitive emergency is usefully thought of as occurring in three dynamically interacting fields: satisfying, specifying and affording. The idea is summarized in Figure 2.

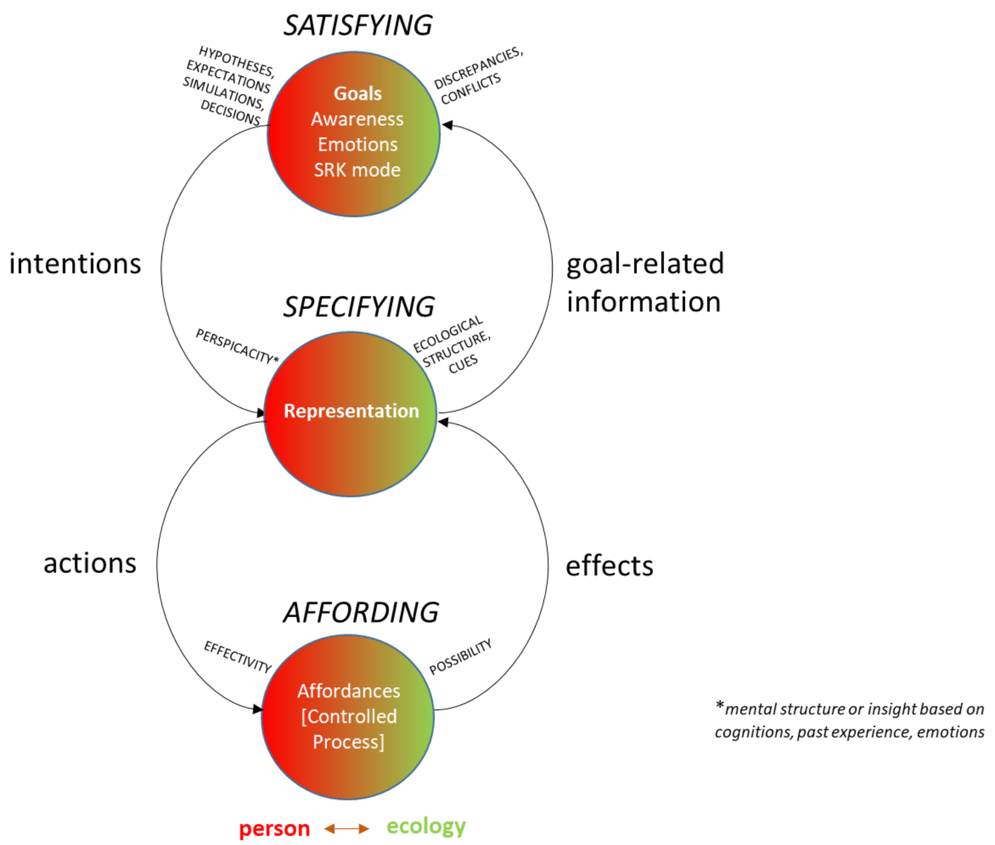


Figure 2. The meaning processing approach to human cognition. Adapted from Flach & Voordhorst, 2020. SRK = Skills-, Rules- or Knowledge-based, after Rasmussen (1983).

According to this approach, the person does not control a process like technology does, with internal algorithms, triggering actuators and responding to sensors (compare Figure 1 and 2). As France (2017) herself noted, human decision making cannot be explained by an internal algorithm. It is affected by emotions, fatigue, and awareness of goal discrepancies, ideas about the past and future, and intentions of other actors. The meaning processing approach is explicit about this, detailing how the person attends and acts based on what matters or *meaning*. It stresses that people often behave *abductively* in complex sociotechnical systems, i.e. approach goals through a process of trial and error in which they evolve strategies by learning from the effects of those strategies on goal discrepancies (costs). Purposeful action (control) evolves the system towards goals on basis of effects of actions on *representations* of the system. We attempt to summarize the key aspects of the meaning processing approach here:

1. In line with control theory, a person seeks to satisfy by achieving goals, evolving hypotheses, expectations and decisions in response to the (potential) costs of goal discrepancies. The process of **satisfying** emerges from the interaction of the person (e.g. physiology, experience, competence, mental and emotional state) and discrepancies apparent from surrounding ecology. Both ecology (e.g. signs, signals) and the person (e.g. fatigue) will influence whether decision-making is skill-, rule- or knowledge-based (Rasmussen, 1983).

1. **Specifying** the goal-related information and how to convert goal-related intentions to actions on the surroundings, involves representations emerging from the interaction between mental and ecological structure, in line with Gibson (2015).
2. Representations are influenced by actions to perceive affordances[1] in the ecology (Gibson, 2015). When conceived of as **affording**, actions on controlled processes depend both on the individual's ability to realize the possibilities that affordances offer, and on the availability of those affordances in the ecology. People perceive meaningful affordances i.e. when operating in-the-loop will not always be mindful about a whole controlled process.

Unlike France's approach, which considers human actions on a controlled process, the meaning processing approach emphasizes that human actions are carried out to satisfy goals, and that in doing so people perceive affordances rather than controlled processes. According to the meaning processing approach, the analyst's job in STPA is to consider how satisfying, specifying and affording affects the controlled process.

AIMS

We aim to assess the feasibility and value of basing STPA-analyses of human actors on the meaning processing approach to cognition. So far, we have:

1) Used STPA to analyse the system of people, technology, infrastructure and procedures working to achieve orderly and efficient contingency parking on terminal closure.
2) Explored how the analysis might be improved by modelling human control as perception and action loops with dynamic coupling of ecological and psychological components, in line with the meaning processing approach to cognition (Flach & Voorhorst, 2020).

We wish to go on and test the feasibility of using meaning processing approach in STPA by using it to structure interviews with actors who decide and act in the study system. This will be reported at a later date.

METHOD

[1] Affordances are aspects of the real ecology that offer possibilities of goal-related action on a system, but they are of no use unless the individual can use them effectively.

STPA analysis was performed on the study system defined under "The real world challenge" using the method of Leveson & Thomas (2018). The analysis was informed by port documents and reports, and three two-hour interviews with two representatives of port management, structured using Leveson & Thomas (2018). To identify possible loss scenarios involving human actors during the analysis, we structured our thinking of human perception and action according to France (2017). Following the analysis, we reflected on how an alternative STPA analysis could be structured to better account for the meaning processing approach to cognition (Flach & Voorhorst, 2020). We also sought to identify and possibly illustrate the potential value of developing STPA using such an approach.

RESULTS of STPA ANALYSIS

The STPA analysis is reported in full in Deliverable 4.2 of the EU project SAFEWAY. To summarise. we identified the following:

- Ten system constraints (SC) that if enforced would help avoid system states that could cause collisions, ship delays or loss of reputation (i.e. system losses). Examples of system constraints are SC1: *Drivers must not drive to a closed terminal instead of car park*; SC2: *No more than X drivers can head to a terminal Y minutes before it closes*.
- Control actions and information flows among the following system components: truck drivers, their employers (transporters), port police, road inspectors, parking staff, port management, on-site coordinators contracted by port management, traffic controllers, and re-routing measures. These are shown in Figure 3.
- 92 unsafe control actions that could violate the ten SCs identified. Unsafe control actions were identified by analyzing how each control action (numbered in brown in Figure 3) could be "unsafe" when carried out, when not carried out, or when carried out too soon or too late. An example of an unsafe control action when control action **14** is carried out is "*Driver tells other drivers via social media to drive to closed terminal*".
- Over 400 scenarios – situations in that would make unsafe control actions likely. In generating loss scenarios, human control was modelled using France's (2017) extension of STPA, i.e. prompts in Table 1 were answered using our knowledge of the system and understanding of truck drivers' situation. Loss scenarios generated for the control action "*Truck driver drives to terminal that is closed or about to close*" (see **13** in Figure 3) are given in Table 2.

To complete the analysis for port management, we designed 29 measures which together would help to avoid all 400+ loss scenarios, maintain system constraints and achieve the aim of port managers. These were derived directly from the loss scenarios e.g. "*Direct driver attention to re-routing measures*" for the first scenario in Table 2 "*Driver does not notice re-routing*".

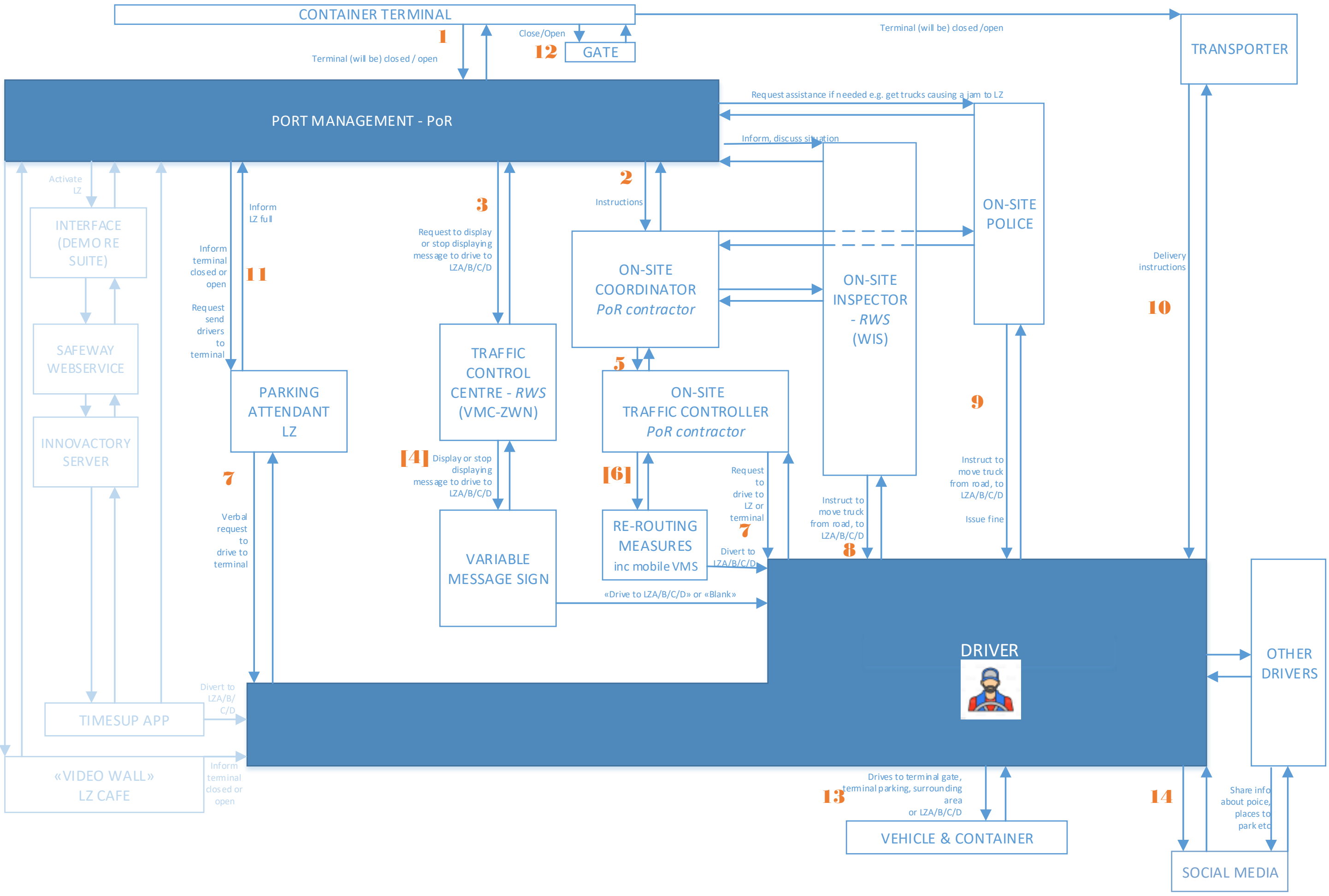


Figure 3. Control structure for the system involved in managing Truck Buffer Scenario. LZ = Landing Zone or holding area; RWS, state road admin; PoR, port management; VMS, variable message sign; WIS, state road inspector, VMC-ZWN regional traffic control center dealing with traffic on roads leading to/from port.

Table 2. Scenarios explaining why drivers could act "unsafely" by driving to a container terminal that is closed or about to close; analysis structured by French (2017)

Driver does not notice re-routing measures or signs
Re-routing measures have not been set out
Driver does not know terminal closed or will close, or when it will close
Driver is confused by conflicting messages from VMS, re-routing measures or Traffic Controllers
Driver attempting to follow re-routing measures to holding area takes wrong turn by accident
Driver believes they can still enter the terminal before it closes
Driver believes terminal is about to re-open
Driver bases action on past experience, when they could queue outside the terminal or park on nearby roads
Driver does not believe that the quickest way to enter the closed terminal is to head for the holding area
Driver heading for another terminal stops at the terminal because of a road blockage
Driver learns from Transporter or Other Drivers that terminal will close but they can enter if they ignore re-routing
Terminal closes suddenly
Driver instructed to head for terminal by Inspectors, Park Attendant or Traffic Controllers
No information given that terminal has closed, will close or when
Inadequate information about need to re-route to holding area or how to do so
No information about negative outcomes of ignoring re-routing at time and place where decision is made
Information about benefits of ignoring re-routing (e.g. jump ahead of queue, save time)
Inconsistent messages from Other Drivers, Transporters, VMS, signs etc. about need to re-routing or fairness of re-route
Driver gets no information on sudden closure when there has been no time to set out re-routing
Driver given incorrect time for terminal closure or terminal needs to change time of closure
Driver gets no feedback that they have taken a wrong turn mistakenly or deliberately
Driver, Transporters get no information about costs of ignoring re-routing measures.

MODELLING HUMAN CONTROL USING THE MEANING PROCESSING APPROACH

In line with the meaning processing approach in Figure 1, we recognized that answers to any one of the three questions in Table 1 also provide answers to the other two questions. For example, what the operator believes about the system affects which action they choose to perform. Reflecting on how unsafe human controller actions could be additionally explained using the meaning processing approach, we therefore identified a single set of prompts to generate answers to a single three-part question;

What is the actor trying to achieve in the system, why (how is the system represented), and how (what affordances are relevant)?

New prompts identified are shown in the left-hand column of Table 3. To illustrate their utility, the new prompts were tested by generating insight related to the single unsafe control action "*Truck driver drives to terminal that is closed or about to close*" (see control action **13** in Figure 3). The results are shown in the right-hand column of Table 3.

The insights are not stated as loss scenarios, but descriptions that *consolidate* and deepen the loss scenarios in Table 2. For example, the first loss scenario in Table 2 is "Driver does not notice re-routing". Table 3 consolidates and adds that this might happen if the operator is in a predominantly "skill-based" mode, e.g. they are operating "on autopilot" because they are familiar with the route (or perhaps have many simultaneous tasks).

The insights in Table 3 also *supplement* the findings in Table 2. For example, they imply that drivers may wish to drive directly to the terminal to merely perceive its status for themselves ("act to perceive"). As part of satisfying, drivers may indeed seek out or attend to any affordance that could allow them to adapt and move closer to their main goal. This could include contacting colleagues or managers to identify workarounds, or continuing on roads leading to terminals to see how close to the terminal they can get. Alternatively, if they have time and are tired, they may choose to stop on way to port to eat or rest, or try to re-arrange delivery plans. They may also try to persuade port or terminal managers to let them enter first on terminal opening.

These and other insights suggest that drivers with different goals and means should be given affordances that bring them closer to their goals. Above all, they should perceive that the holding area will afford the earliest entry to re-open container terminal. In this regard, signal detection is a central aspect of control theory (Feufel and Flach, 2019), and in answering the new prompts we see that the salience of information allowing drivers to distinguish whether (i) diverting or (ii) heading directly for the closed terminal is most likely to afford punctual delivery is paramount.

Insights like these give a more in-depth understanding of how human actor cognitions in real situations might promote or inhibit unsafe control actions. In this way the meaning processing approach can help close the gap between identification of unsafe control actions and design of measures to help avoid them.

Table 3. Prompts analysts can use to consider scenarios explaining why humans could act unsafely (left column); illustrated with information triggered by considering the action "Truck driver drives to terminal that is closed or about to close".

Prompt	Why drivers would drive to closed terminal: Insight
How does satisfying affect ecological search and representation of affordances?	Drivers have delivery deadlines and activities planned and emerging before and after delivery; they will attempt to reduce goal discrepancy by actively seeking out ways to deliver as quickly as possible. They may not attend to re-routing signs because: they are operating in skill-based mode (fail to observe road signs as driven this way many times before); they "act to perceive" the terminal entrance (main affordance) for themselves; re-routing signs have been wrong in the past; they believe terminal managers will let their delivery in; they are not aware of costs or negative affordances e.g. they will increase delays for themselves and others.
How does specifying affect: (i) availability and use of affordances? (ii) goal-related information?	In absence of other goals, drivers will search for affordances that allow them to satisfy delivery goals e.g. open roads to terminal entrance, messages from other drivers (saying terminal still open, possible to wait on roads), heavy traffic at holding area, difficulties in turning large vehicle (effectivity). They may oversee or miss affordances that allow them to satisfy goals they are not interested in satisfying e.g. services.
How does affording influence representation of the system and goal achievement?	As long as use of open roads leading to terminals aligns with perspicacity and expected ecological structure, the system will be represented as one in which is bringing driver is closer to goal (positive effects of affordances).
Are there several goals the person wants to satisfy? How are competing goal discrepancies traded off?	Safe driving, pleasing managers, resting or refreshing, sleeping, need to get home etc.? Drivers may also need to deliver to comply with driving hours legislation.
Which goals and goal discrepancies is the person aware of (or not aware of)? (Includes consideration of designer vs. operator goals.)	Port management may assume driver goals are to deliver in compliance with diversions.
How does the person's emotional or physiological state affect how goals are attended to?	Important to consider that drivers may be in need of rest, or concerned about knock-on effects of not delivering.
Are decisions about intentions predominantly: skill-based (was a familiar action attempted when not appropriate?); rule-based (which rules were applied?); or knowledge-based (what mental model of system or process?)	Driver may deliver in skill-based mode, not attending to diversions; driver may believe (from others or past experience) that the quickest way to deliver is to ignore diversions (rules); or may simulate way to deal with scenarios in which they are turned round for not complying with re-routing signs (knowledge-based).
As they attempt to move closer to goals, what affordances would the operator want to try out?	Contact colleagues and managers to inform or be informed; try out open roads leading to terminals; stop on way to port to eat or rest; try to re-arrange delivery plans; try to persuade port managers or terminal to let them enter first on terminal opening.
What are the costs of: (i) Suboptimal use of affordances? (ii) Use of goal-mismatched affordances? (iii) Use of poorly designed affordances?	The driver may not see any costs of using open roads leading to closed terminals; may perceive and/or experience costs of diverting to holding area.
How can personal and ecological factors (interact to) affect representation (e.g. physiological factors, experience, "ecological structure", situational complexity, physical conditions)	Driver does not understand re-routing signs; driver may not be familiar with holding area and not want to use it; driver may be overloaded; poor weather may cause delays making driver think terminal may have re-opened.
What goal-related signals are salient?	Not clear whether information about costs or benefits of diverting most salient for drivers. Related to the container delivery goal, the most effective "field of safe travel" on road to closed terminal will be salient.
Which process mindsets are likely?	If I drive a little faster, I could enter the terminal before closure The quickest way to enter re-open terminal is to re-route vs. ignore diversions
What ideas are there about intentions and actions of other people and technology?	Intentions of colleagues (other drivers may try to jump the queue) Intentions of managers, police, inspectors (Port managers don't care if we wait outside the closed terminal) Intentions of employer

DISCUSSION

Preliminary analysis suggests that supplementing France's (2017) STPA extension with Flach & Voorhorst's (2020) meaning processing approach to human cognition can deepen understanding and give new insight into scenarios explaining why humans would act "unsafely" potentially causing system loss. Not surprisingly, given common methodological foundations, such as skills-rules-knowledge modes of cognition (Rasmussen, 1983), the Recognition-Primed Decision Making model (Klein, 2017), and control theory (Jagacinski and Flach, 2009), there is overlap between the prompts and questions the analyst should consider using France's (2017) "human extension" to STPA and the Flach & Voorhorst's (2020) meaning processing approach to human cognition. Nevertheless, the meaning process approach causes the analyst to deliberate on additional important aspects of control that are unique to humans. For example, it highlights that human "controllers":

- Will engage in dynamic control of the situation to stay as close as possible to goal achievement (e.g. deliver a container), and in doing so may act to perceive (e.g. drive to a closed terminal to see if it has opened) as well as perceive to act (e.g. see signs and follow diversions)
- Do not perceive a "controlled process" but the meaning (goal relevance) and value (positive or negative) of affordances
- Perception and action depend on how the surrounding ecology is represented; perception and action emerge from interactions with the ecology
- The use and effect of affordances depends on whether they capture the awareness of the person (depends both on ecological structure and on meaning and value of affordance to the person) and how effectively the person can use them
- The person will rarely perceive data from sensors alone, but interpret a dynamic array of information in the ecology, which they will relate to one or more goals
- The person will often have several personal and professional goals, and action can move the person towards (gains) or away (costs) from each goal

Once unsafe control actions are identified using STPA, the meaning processing approach can bring the analyst closer to design solutions, because of its emphasis on ecological psychology. For example, by asking about which affordance alternatives are available to the human controller, rather than which action alternatives, ideas are generated that have direct relevance for design during the analysis.

While the insights we have generated in this article illustrate the approach, they only address the control-related cognitions and actions of a single actor (the truck driver). But the meaning processing approach could be applied to several human actors in a system – preferably using knowledge elicitation techniques from naturalistic decision-making – such that problems that emerge from interactions among actors can be understood. We hope to try this out in future projects, in which we will also test the new prompts using knowledge elicitation techniques with drivers or other human controllers.

One drawback of extending the STPA approach as we describe, however, is the extra resource and understanding required. Indeed, an important aim of France's (2017) model was to provide engineers with a simple account of human control that would not require in-depth understanding of psychology; and the use of the meaning processing approach would not seem in line with this aim. While it may be possible to simplify the meaning processing approach, we ask whether engineers should attempt to understand people using a superficial approach, given that people are inordinately more complex and difficult to understand than the technology used in contemporary systems.

IMPACT

Based on preliminary analyses, we believe that the meaning processing approach to cognition could be used to obtain a less piecemeal understanding of driver behavior, one that leads to more complete and satisfying sociotechnical solutions to traffic management and other control challenges. The meaning processing approach can be used to develop solutions that better account for adaptive human control, and to recommend supplementary analyses for STPA practitioners.

REFERENCES

Flach, J., Voorhorst, F., 2020. A meaning processing approach to cognition, Resources for ecological psychology. Routledge.

France, M., 2017. Engineering for Humans: A New Extension to STPA (Masters). Massachusetts Institute of Technology, Department of Aeronautics and Astronautics.

Gibson, J., J., 2015. The ecological approach to visual perception, Classic Edition. ed. Psychology Press, Taylor & Francis, New York.

Jagacinski, R., J., Flach, J.M., 2009. Control theory for humans. Quantitative approaches to modelling performance. CRC Press, Boca Raton, FL.

Klein, G., 2017. Sources of Power: How People Make Decisions, 20th Anniversary Edition. ed. MIT Press, Cambridge MA.

Leveson, N., 2012. Engineering a safer world: Systems engineering applied to safety, Engineering Systems. MIT Press, Cambridge MA, London England.

Leveson, N., Thomas, J., 2018. STPA Handbook.

Montes, D.R., 2016. Using STPA to inform Developmental Product Testing (PhD thesis). MIT, Boston, MA.

Rasmussen, J., 1983. Skills, Rules, and Knowledge: Signals, Signs and Symols and Other Distinctions in Human Performance Models. IEEE Transactions on Systems, Man and Cybernetics SMC-13.

Thornberry, C., 2012. Extending the human-controller methodology in systems-tehoretic process analayses (Masters Thesis). MIT.

Vicente, K., 1999. Cognitive Work Analysis. Toward safe, productive and healthy computer-based work. LEA, New Jersey.

Funding: The SAFEWAY project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 769255. This paper reflects only the views of the authors. Neither the Innovation and Networks Executive Agency (INEA) nor the European Commission is in any way responsible for any use that may be made of the information it contains

Highlights of Recent Work

Operational Resilience Management

Daniel LICHTE[a], Evelin ENGLER[b], Frank SILL TORRES[c] and Eva BRUCHERSEIFER[a]

[a]*German Aerospace Center, Institute for the Protection of Terrestrial Infrastructures, Germany*
[b]*German Aerospace Center, Institute of Communications and Navigation, Germany*
[c]*German Aerospace Center, Institute for the Protection of Maritime Infrastructures, Germany*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

- We elaborate how the coordination of resilience management means and risk mitigation measures can be executed with the help of enhanced situational awareness.
- We discuss how the concept "Digital Twin" can contribute to an innovative and operational resilience management for socio-technical systems and organizations.

2. Summarize New results

This work is a theoretical study on the applicability of the concept "Digital Twin" in the context of resilience management framework. Initially, main requirements on such a framework have been studied and structured. Based on that, the concept of the "Data and Information Base" (DIB) has been developed. Such a DIB refers to a structure that hosts the data of the system and environmental models, information about the employed sensor system as well as the acquired sensor data, and the so-called scenario spaces, which represent possible scenarios that may be described by the available system and environmental parameters. A further result is a detailed process model for an operative resilience management that integrates the DIB. It is a notable observation that the inherent abilities of real-time situational analysis with feedback to the simulation of situation developments may lead to enhanced situational awareness. Furthermore, the extensive simulation of all feasible scenario developments and measure effectiveness may enable enhanced decision-support.

Despite its level of detail, the developed framework remains on a higher level of abstraction. Future works will focus on the integration and experimentation of the framework in real-world applications.

3. Who would benefit from knowing about your work

This work is mainly directed to decision-makers and researchers in the field of risk and resilience management.

4. The impacts or implications of the results or how they can be utilized

The results of this work might pave the way for a risk and resilience management that profits from the advances in the fields of computer science and digitalization. The application of the concept "Digital Twin" in the context of risk and resilience management is rather new, but offers promising opportunities that can lead to more resilient socio-technical systems and organizations. The framework was designed with the requirement of operability, i.e. we put high emphasis on the practicability of the tasks and processes.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Increasing complexity and interdependencies are prominent features of current organizations and technical or socio-technical systems while digitalization is further accelerating this process. At the same time, the goal of withstanding adverse events remains unchallenged. However, classic risk management processes are reaching its limits when coping with such complex structures. This led to rising interest in the resilience concept, which extends the risk centered approach by a strong focus on inherent capabilities of systems or organizations. Several recent works focused on the definition of the concept of resilience. Though, only few advancements were made towards the actual implementation and operation of the resilience concept as management process in organizations or systems. Based on this motivation, a framework for an operative resilience management for socio-technical system or organization (STSO) has been elaborated. This framework describes a process that administrates and improves the resilience of STSO by providing monitoring, responding, anticipation and learning capabilities. Furthermore, the situation-driven approach is able to enhance these capabilities. Additionally, digital concepts, especially the concept of digital twins, are deeply integrated into the framework via a data and information base (DIB) that contains models and information about the monitoring system, the environment and scenario spaces. This work presents and discusses in detail the DIB and the tasks of the resilience management process, while focusing on operable solutions.

It can be shown at hand of the proposed framework that an operational resilience management is a complex task with promising abilities, especially when implemented using advanced concepts like the digital twin. It is a notable observation that the inherent abilities of real-time situational analysis with feedback to the simulation of situation developments may lead to enhanced situational awareness. Furthermore, the extensive simulation of all feasible scenario developments and measure effectiveness may enable enhanced decision-support. This is further strengthened by the inclusion of risk-benefit ration of measures, resulting uncertainties in addition to measure efficiency.

Highlights of Recent Work
Learning to Adapt, as we Adapt: A Hospital's Experience of Learning During the Pandemic

Sudeep Hegde[a], PhD, Ethan Larsen[b], PhD, Siddarth Ponnala[b], PhD, Erin Pohl[b], MS, Raymond Sze[b], MD, Monica Miranda Schaeubinger[b], MD, MSPH
[a]*Clemson University, United States*
[b] *Children's Hospital of Philadelphia, United States*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

- Our work demonstrates a learning tool and insights about on how a large hospital adapted its processes to meet the exigent challenge of the pandemic.
- The study has identified multiple instantiations of adaptive capacity and organizational synergy developed during this crisis – capabilities that can be sustained beyond the pandemic itself.

2. Summarize New results

Previous work exists on the use of knowledge-elicitation techniques to proactively learn about successful adaptative practices. In particular, the Resilience Engineering Tool to Improve Patient Safety (RETIPS) has been used on a pilot-basis to learn about adaptive capabilities of anesthesia residents3. RETIPS is a semi-structured self-reporting tool designed to elicit narratives of adaptation in everyday clinical work. The current study tailored RETIPS to the case of a radiology unit to learn from its operational staff about what went well as they adapted to the pandemic.

The version of RETIPS used for this study includes the following questions:

- Role and years of experience
- Detailed narrative: example of an adjustment or coping strategy or a preventive measure. Cognitive probes include types and sources of information shared, anticipation, innovation, infection control measures, monitoring behaviors, resource availability, policies and procedures, cooperation with colleagues, and engagement from leadership.
- Multiple choice questions (checkboxes) in the following categories: What is Working Well; Challenges and Concerns; Resources; Area of Practice; Surprises; and Discovered Opportunities.

The tool was made available via Google Forms to all employees of the hospital on July 29 2020. The majority of responses were received over the course of 10 weeks, after which the analysis of data began. The narrative responses were analyzed for qualitative themes relating factors that enabled success, including strategies and tactics described in the examples.

Fifty-eight reports were received between July 29th and October 12th 2020 with over 90% received during the month of August. Participants' roles included radiology leadership (21%), technologist (19%), nursing (13%), radiologists (10%), and sedation/anesthesia (10%). Participants described examples of lessons-learned, and indicated the success factors, challenges, and resources pertinent to their examples. The thematic analysis of the descriptions of their examples of an 'adjustment or coping strategy or preventive practice' identified several prominent themes, briefly described below:

- *Communication of change:* related to mention of information dissemination mechanisms for policy or procedural change. These included communication via email, town halls, huddles, and other forms of briefing.
- *Knowing where to find resources:* indication that information or resources were sought out and retrieved rather than having to wait for them to be pushed to the participant. Participants cited availability of resources as well as information about access as important factors for success. Key resources included personal protective equipment (PPE), and cleaning materials and sanitizers.
- *Culture and attitudes:* indication of the nature and mindset of individuals the respondent encountered in their narrative.
- *Timeliness and clarity of information:* mention of the time frame information from leadership or higher is provided or the clarity of the message.
- *Leadership:* direct mention of supervisor or beyond roles relative to the respondent.
- *Cooperation between coworkers:* indication that shared workload, or collaboration is occurring. Respondents indicated a high level of cooperation and sharing between colleagues and coworkers.

- *Interdepartmental cooperation:* Coordination and sharing of resources between various departments within the hospital.
- *Shared understanding and awareness:* shared mental models and expectations among coworkers and teams.
- *Accountability:* Sense of ownership and responsibility in individual actions.
- *Availability of policies and procedures* to address a variety of situations.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Organizations and other multiscale systems that have grappled with the pandemic or a similar crisis and are looking to learn from their own experiences will benefit from this work. The tool RETIPS can be adapted to the specific domain-type and needs of the organization and used as part of its overall learning process.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The study demonstrated the feasibility of implementing RETIPS as a tool for the organization to learn from individual experiences at the frontlines of care. Several of the themes described above, represent capabilities that were either discovered, re-discovered, developed or enhanced during this extended crisis. A large part of the hospital's success in adapting can be attributed to the frequency and timeliness of communication between the senior leadership and frontline workers. An active synergy has been developed and maintained between the blunt- and sharp-ends of the organization. The lessons from this experience need to be learned in order to sustain these capabilities going forward, even beyond the pandemic itself.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The COVID-19 pandemic is a case of fundamental surprise, a crisis that emerged suddenly, without warning, and which had no preceding pattern. Like most organizations, hospitals were caught off-guard with few to no pre-existing protocols, norms, standards or guidelines in place. Furthermore, public health and institutional leadership did not have enough available information to make informed decisions in the beginning. Therefore, many time-sensitive decisions have had to be made under ambiguity and uncertainty. Hospitals' management and operations have been challenged with elevated patient volumes; infection prevention; lack of resources; decisions about mode of care-delivery. Hospitals have been hit with a paradox: while it's increasingly important to safely deliver care, it's more difficult to do so. During the past few months, however, order has emerged from the chaos even as the pandemic continues. Our work represents a proactive learning mechanism to extract insights in real time from an organization's own adaptive efforts. As such, the learning effort and the insights generated could be extended as part of an organization's preparedness of fundamental surprise events in general.

References

1. Romani G, Dal Mas F, Massaro M, et al. Population Health Strategies to Support Hospital and Intensive Care Unit Resiliency During the COVID-19 Pandemic: The Italian Experience. *Popul Health Manag*. December 2020. doi:10.1089/pop.2020.0255
2. Enabling Health Care Resilience During and After the COVID-19 Pandemic. https://healthpolicy.duke.edu/news/enabling-health-care-resilience-during-and-after-covid-19-pandemic. Accessed February 25, 2021.
3. Hegde S, Hettinger AZ, Fairbanks RJ, et al. Qualitative findings from a pilot stage implementation of a novel organizational learning tool toward operationalizing the Safety-II paradigm in health care. Appl Ergon. 2020;82:102913. doi:10.1016/j.apergo.2019.102913

Highlights of Recent Work

Supporting Resilience in Large-scale Multi-domain Operations as Learning Laboratories

Martin VOSHELL[a], James TITTLE[b], Tod FINGAL[a], Robert RICKARD[a]
[a]Rickard Consulting Group, Inc
[a]MISI DreamPort

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- For military agencies worldwide, adversary threat capabilities continuously place new demands on weapon system training and mission and package employment that reveal critical integration challenges and new contexts for expertise and continuous learning.
- By looking at recent large-scale multi-domain air warfare exercises through the lens of the learning laboratory framework, this paper describes how next-generation data integration of multiple domains must be used to study and develop analytic insights into modeling resilience in complex mission environments especially in the context of integrated Live, Virtual, Constructive (LVC) training events.

2. SUMMARIZE NEW RESULTS

Abstracting research from large-scale field observations and exercises can be a daunting challenge for researchers and practitioners. Voshell and Woods (2008) developed a conceptual framework of the "learning laboratory" as a cognitive systems engineering (CSE)-based research and design approach to abstract valuable lessons learned and sharable knowledge from staged world studies and training exercises. The original learning laboratory concept presented a high fidelity yet controlled environment to study team processes, technology applications, and system resilience in rare and critical event contexts. Such constructs have been successfully used in the past to examine decision making and system resilience in emergency response exercises (Voshell et al, 2008), high fidelity nuclear power plant simulations (Roth, Woods & Pople, 1992), and human-robot teaming for search and rescue (Murphy, Burke & Stover, 2006). Conceptually, a learning laboratory is an exercise based on CSE scaled-world principles that results in local, sharable, and generalizable learning which contributes to the understanding of cognitive work and system resilience in a work domain. To support discovery and effective learning in large scale exercises, the learning laboratory approach uniquely focused on planning for learning opportunities and anticipating scaling challenges from multiple dimensions. This approach enables multiple stakeholders to plan, collect data, and evaluate performance trends to support a variety of abstract goals and learning purposes. In a learning lab, one of the most critical components to instruction and learning is the debrief or after action review (AAR), where the art of figuring out "what" happened and then "why" triggers reflection (as lessons learned or debrief focus points (DFPs)). As outlined by Wiese, Freeman, Salter, Stelzer, and Jackson (2008), the debrief should function to make sure the right participants learned the right lessons so that they will recall correct and incorrect performance and the context around it so that learning transfer is appropriately encouraged or discouraged in future situations. Effective debriefs help participants discover performance failures and successes, reflect on them, and diagnose causes. The debrief should support social evaluation of performance and competence across other participants. What debriefs do not do, however, is always abstract to generalizable learning. While this process can be very effective for local and sharable learning, it is this further abstraction toward generalizable learning at an understanding level that is the most critical to realize the full benefits of effective training. From this conceptual lens, the authors have been able to reflect upon their work supporting modern large-scale multi-domain operation (MDO) military training exercises as a natural laboratory to explore cognition at work and system resilience.

To support rapid learning and operational improvement at scale, the US Air Force Warfare center conducts one of the largest training exercises in the world known as "Red Flag". Developed in the Korean war to provide realistic combat training through the use of complex mission scenarios, Red Flag has evolved to include joint service participants from the United States and over 35 allied nations' combat Air Forces, in Air, Space, and Cyber domains. From a learning lab perspective, training stakeholders at such events face a prediction role, requiring them to generate hypotheses about future operations, challenges, new threat capabilities, and tactical employment advances. Three times a year, training stakeholders must continuously incorporate new technologies to help the operational community train against enemy advanced threats and scenarios, with numerous friendly technologies and tactics complicating the planning. Stakeholders work closely with the operational community to combine live flight operations with virtual simulations and computer-generated forces that real-world live assets cannot physically replicate in a useful manner to maximize the usefulness of expensive and time-consuming live training exercises. However, fully supporting sharable learning at such exercises remains a continual and critical support function and is considerably hampered for a single reason – the data from each environment (Live, Simulated/Virtual, and Constructive) is incongruent and difficult to homogenize.

One of the most enduring organizational process constructs in Air Force fighter communities is the "debrief culture" (Causey, 2020). Standard debrief templates are structured around identifying desired learning objectives (DLOs), which represent goals which are then identified, decomposed, and focused on during analysis and review. DLO success and failure take the form of debrief focus points (DFP) and learning points (LPs), which represent the key information relationships in a debrief to frame key lessons for exercise participants. Despite the central role of the debrief in training exercises, software support tools have failed to directly support key decision-making requirements critical for effective debrief support due to the incongruency of the data sets. As a response to these challenges the authors have fielded advanced debrief tool concepts from a learning laboratory approach incorporating design patterns from resilience engineering as part of a multi-year research and development effort. Developed closely with warfighters in advanced training exercises, the team was able to capture and facilitate sharable learning and develop analytic insights that complemented debrief culture. Using this debrief support capability in a large-scale exercise led to the following capabilities and insights:

- DFPs and LPs, framed by LOs are captured, time-stamped, and saved along with full-mission data recordings augmented by advanced visualizations, and multi-dimensional analytics to capture full joint system performance.
- Using the novel debrief tool, subjective human performance was coupled with objective data gathered from live missions for the first multi-domain/objective/subjective data collation events. This uniquely integrated warfighters from space, cyber, and air domains who were represented equally in after action review, providing true MDO lessons learned.
- While developed as a mission training tool- the developed system has great potential as an R&D tool to explore resilience in natural laboratories. While recent work has looked at formally modeling air power operational resilience, for example Hagen et al.'s (2016) framework based on sequential game theory, the same modeling and simulation functions can now be implemented for experimentation against tactical reality to assess real-time live mission performance.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

Researchers and developers of training systems and debriefing systems, as well as problem holders and organizational training stakeholders interested in incorporating technologies in the support of improved team training. This work is also applicable to researchers studying the impact of Live, Virtual, and Constructive (LVC) environments on team training effectiveness from the perspective of focusing on sharable learning

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

Debrief tools developed using learning laboratory design concepts can be highly useful for both operational training as well as for conducting advanced system engineering research and development to leverage the traceability available in modern instrumented exercises to explore and analyze complex system adaptability, extensibility, and operational resilience.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Training exercises across domains are often conducted to help better anticipate and accommodate new technology integration and new tactics and skill development. This requires an understanding of the work and frames researcher and engineer responsibility for staging exercises effectively as learning laboratories to steer learning opportunities in such a way as to get useful learning for all involved participants and ultimately facilitate insight into new understanding about the nature of work in the domain. For broader applicability, the same design approach and toolset can be adapted to evaluate performance in real-world operations as part of an overarching process of learning.

PLEASE LIMIT YOUR SUBMISSION TO TWO PAGES OR LESS

Highlights of Recent Work

Training for Macrocognitive Skills Awareness in Cybersecurity Professionals

HIFSA A. KHAN, MICHAEL R. HELZER, DAVID SCHUSTER
San José State University

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

In this presentation, we will discuss the implications of recently-identified macrocognitive themes relevant to cybersecurity defenders (Schuster, in press). Further, we will propose a novel training strategy and research study needed to understand how to effectively train macrocognitive skills.

2. Summarize New results

As the prevalence of security breaches increases, cybersecurity is becoming increasingly important in sectors such as healthcare, government and business. (Hemphill & Longstreet, 2016). Cyber-defenders in particular carry out daunting and critical work to defend information systems from ever-increasing vulnerabilities (Shah et al., 2018). While these developments have prompted an interest for increased recruitment and training (Brannen, 2013), currently very little research has emphasized the role of cognition in cybersecurity training and job performance.

In our forthcoming publication (Schuster, in press) we identified seven macrocognitive themes resulting from knowledge elicitation interviews with cybersecurity professionals. These macrocognitive themes describe how cognitive processes support the proficiency of cyber defenders and are a starting point for describing cognition specific to cyber defense. In our presentation, we will present each of these themes and describe their implications for research and practice.

Given the urgency and magnitude of the cybersecurity problem, there is a great need to accelerate the development of cybersecurity professionals (Brannen, 2013; Crumpler & Lewis, 2019). To meet this need, we describe a method of training macrocognitive skills in cyber defenders and propose a study to evaluate its effectiveness. Specifically, we propose awareness training for macrocognitive skills. Taking inspiration from crew resource management (Wiener et al., 2010), which teaches teamwork to flight crews as the use of “all available resources in the cockpit” (Lauber 1984, p. 20), our training would present macrocognition as key to leveraging all available cognitive resources in the security operations center. Essentially, we propose training to facilitate metacognition, which is defined as individuals’ knowledge and regulation of cognitive processes such as memory, attention, and reasoning (Forrest-Pressley, MacKinnon, & Waller, 1985; Schraw & Dennison, 1994). Past research has suggested that metacognition is linked to improved pattern recognition and performance in novel and uncertain contexts (Anderson et al., 2006; Cohen et al., 1993; Cohen, Freeman & Thompson, 1995). This may be because metacognition promotes self-regulation and iteration in the problem-solving process (Anderson, 2006; Efklides, 2006). This improved performance following metacognitive training exemplifies that metacognition is an important skill of experts (Cowen, 2014; Hoffman & Woods 2011). If we can augment cybersecurity training by including metacognition elements, novice cybersecurity defenders may acquire the soft skills of experts in a quicker fashion than the traditional means.

The goal of the current study is to create and evaluate a simple awareness training that introduces macrocognitive and metacognitive concepts to novice cybersecurity defenders to improve their performance and retention. Our research questions are: Does awareness and understanding of macrocognitive concepts increase with training? Does defender performance improve after learning macro cognitive concepts? What connections do defenders draw between macrocognitive concepts and on-the-job tasks?

First, we identify the macrocognitive concepts of interest, using (Schuster, in press) as a guide and focusing on concepts that have the most support in the literature for cybersecurity and related fields. Next, we will define the skills and knowledge to be trained. We will illustrate our approach with an example. Finally, we will describe how we will evaluate our training at the reaction and learning levels defined by Kirkpatrick (1959, 1996). The reaction-level responses may provide insight into what macrocognitive concepts are relevant to cybersecurity professionals after they are familiar with the concepts. This can inform research in cybersecurity training for macrocognition.

The proposed study serves two aims. First, it could provide evidence for the efficacy of metacognitive training for macrocognitive skills. Second, a distinguishing feature of the proposed study is the capture of qualitative feedback from subject-matter expert (SME) participants. Feedback from the training can increase understanding of the macrocognitive skills involved in cyber defense.

3. Who would benefit from knowing about your work

Macrocognition researchers, cybersecurity researchers and entities which rely on the services of cybersecurity defenders would benefit from this study. Macrocognition researchers will benefit from this study because it aims to apply macrocognitive concepts to complex and uncertain cybersecurity contexts beyond the conceptual level. Although several task analyses have described macrocognitive themes in cybersecurity (e.g., Paul & Whitely, 2013), few are incorporating these ideas into training. This proposed study will also be beneficial to cybersecurity researchers because it provides empirical evidence for a novel training method. Adding to the training literature would allow improved training programmes and hence more opportunities to improve the performance of cybersecurity practitioners. Finally, any business or government agency who is reliant on cybersecurity professionals could benefit from this work, because of the potential for the development and implementation of macrocognitive training to improve job performance for cyber defenders.

4. The impacts or implications of the results or how they can be utilized

Researchers will benefit from a review of existing literature to identify cognitive skills relevant to the roles of a cybersecurity professional along with an empirical foundation for the training of such skills. We propose that future studies could utilize SMEs to develop task-specific situations which could be used to test the efficacy of cognitive skills training on improving the job performance of new cybersecurity defenders.

5. The wider scope or relevance of the work

While cybersecurity experts are in high demand, there is not a clear educational path for emerging professionals to fill the gap in the market (Crumpler & Lewis, 2019). Furthermore, although research is advancing the state-of-the-art in the automation of network defense, little research has been conducted on the role of human cognition in the success of cybersecurity defenders, which remains critical in this complex environment. As a result of this, early-career cybersecurity defenders may not be well prepared for their roles initially (Crumpler & Lewis, 2019). If cognitive skills that directly contribute to performance in cyber roles are identified and trained, cybersecurity professionals may be trained faster and more efficiently.

Acknowledgement

This work is based upon work supported by the National Science Foundation under Grant No. 1553018. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Anderson, M., Oates, T., Chong, W., & Perlis, D. (2006). The metacognitive loop I: enhancing reinforcement learning with metacognitive monitoring and control for improved perturbation tolerance. *Journal of Experimental and Theoretical Artificial Intelligence* 18(3), 387–411.

Brannen, K. (2013, January 31). *Wanted: Geeks to help fight Pentagon's cyberwar.* Politico.com. Retrieved February 11, 2013, from http://www.politico.com/story/2013/01/wanted-geeks-to-help-fight-pentagons-cyberwar-86946.html

Cohen, M. S., Adelman. L., Tolcott. M. A., Bresnick. T. A., & Marvin. F. F. (1993). *A cognitive framework for battlefield commanders' situation assessment* (Tech. Report 93-1). Arlington, VA; Cognitive Technologies, Inc.

Cohen, M. S., Freeman, J. T., & Thompson, B.B. (1995). *Training the naturalistic decision maker* (Tech. Rep. under contract MDA903-92-C-0053 for the Army Research Institute). Arlington, VA: Cognitive Technologies.

Crumpler, W., & Lewis, J. A. (2019). *The Cybersecurity Workforce Gap*. Center for Strategic and International Studies. https://www.csis.org/analysis/cybersecurity-workforce-gap

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, *1*, 3–14.

Forrest-Pressley, D. L., MacKinnon, G. E., & Waller, T. G., (1985). *Metacognition, cognition, and human performance*. Academic Press.

Hemphill, T. A., & Longstreet, P. (2016). Financial data breaches in the U.S. Retail economy: Restoring confidence in information technology security standards. *Technology in Society, 44*, 30-38. https://doi.og/10.1016/j.techsoc.2015.11.007

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training Directors*, *13*, 3–9.

Kirkpatrick, D. L. (1996). Invited reaction: reaction to Holton article. *Human Resource Development Quarterly, 7*, 23–25.

Lauber, J. K. (1984). Resource management in the cockpit. *Air Line Pilot*, *53*, 20– 23.

Paul, C. L., & Whitley, K. (2013). A Taxonomy of Cyber Awareness Questions for the User-Centered Design of Cyber Situation Awareness. In L. Marinos & I. Askoxylakis (Eds.), *Human Aspects of Information Security, Privacy, and Trust* (8030, pp. 145–154). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39345-7

Schraw, G., & Dennison, R. S. (1994). Assessing Metacognitive Awareness. Contemporary Educational Psychology, 19(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Shah, A., Ganesan, R., Jajodia, S., & Cam, H. (2018). A methodology to measure and monitor level of operational effectiveness of a CSOC. *International Journal of Information Security, 17*(2), 121–134. https://doi.org/10.1007/s10207-017-0365-1

Wiener, E. L., Kanki, B. G., Helmreich, R. L., & Anca, J. M. J. (Eds.). (2010). *Crew resource management* (2nd. Ed.). Academic Press.

Highlights of Recent Work

How can we help novice child protection social workers to see situations like experienced practitioners? A randomised controlled trial evaluation of the ShadowBox™ method using pre-recorded video feedback

WHITTAKER, A.[a], MOSER, K. S.[a,b], and HARVEY, D.[a]

[a]London South Bank University, UK
[b]University of Queensland, Australia

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

- Many organisations, such as child protective services, need experienced practitioners but face real challenges in staff retention.
- The ShadowBox™ method enables the large-scale online training of novices using pre-recorded video feedback and this paper outlines the promising results of an RCT evaluation.

2. SUMMARIZE NEW RESULTS

Background

Protecting children for abuse and neglect is a complex area of decision-making but frequent staff turnover has meant that many frontline child protection social workers are often relatively inexperienced. The ShadoxBox™ method is an educational intervention that enables novice practitioners to gain decision-making skills quickly.

Aim:

This study was an RCT evaluation of an educational intervention for novice social workers that used the ShadoxBox™ method adapted to include pre-recorded video feedback from an expert panel to test whether novice decision making would become more similar to experienced practitioners.

Methods:

The study was an RCT in which participants completed complex scenarios in a computer lab in control and intervention groups. The training method involved participants receiving feedback from highly experienced practitioners at each decision point within the scenarios. The intervention group received pre-recorded video feedback from a panel of highly experienced practitioners while the control group received no feedback.

Participants and Setting

Participants (n=83) were trainee social workers from a London university randomly allocated to control and intervention groups. Data was collected on computer stations using Qualtrics.

Results

Undertaking the scenarios improved both groups but greater improvement was seen in participants who had received video feedback from a panel of experienced practitioners. The results were promising, with participant accuracy increasing by 44% (from 31% to 75%) in the intervention group compared to an increase of 31% (from 32% to 63%) in the control group.

Considerable improvements were noted in both intervention and control groups, which suggest that scenario-based interventions can be a promising educational method as learning is rooted in real life scenarios and participants have the opportunity to reflect upon their decisions.

The qualitative findings are that novice participants make predictable errors, including:

- Making shallow assumptions
- Focusing exclusively on the parents rather than the child.
- Jumping to early conclusions with insufficient information.

Conclusions

- ShadoxBox™ training appears to be a promising intervention for improving decision making.
- Novices benefitted from having concentrated exposure to complex scenarios focused on assessing risk and making professional judgements.
- When this was augmented by direct feedback from a panel of highly experienced practitioners, these benefits were increased considerably.
- The complexity of the scenarios also exposed the novices to real life pressures rather than the simplified versions used in decision research.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK?

- NDM researchers.
- Manager of organisations that want novice staff to gain the understanding of experienced staff in a time and resource-effective way.
- Designers of online training.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

- The study is a development of the ShadoxBox™ model to incorporate video feedback.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

- The findings support the value of scenario-based training as improvements were found in both groups.
- Greater improvement was seen in participants who had received expert video feedback, which supports the value of the ShadoxBox™ method.

Acknowledgements: The researchers would like to acknowledge the support of the British Academy and Leverhulme Trust (grant number SG160251).

A Systems-Resilience Approach to Technology Transition in High-Consequence Work Systems

Kelly NEVILLE, Heather ROSSO, and Bianica PIRES
The MITRE Corporation

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

Our main objective is to produce a paradigm shift in the way technologies are transitioned into high-consequence work systems. The new paradigm is a systems-resilience approach to Information Technology (IT) transition that will improve the rate at which technologies successfully transition and do so safely, effectively, and without compromising the receiving system's ability to perform and adapt in high-demand conditions.

2. NEW RESULTS

Our team will present an initial version of our Transform with Resilience during Upgrades to Socio-Technical Systems (TRUSTS) Framework, which specifies sources of system resilience that need to be preserved during technology transitions and system modernizations. We seek community discussion of the framework and early ideas for framework-based assessment tools that will be used to guide transition and modernization.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

This work focuses on improving the IT transition and integration of new technology into high-consequence work systems such as surgical and other healthcare delivery systems, military operations, transportation systems such as air traffic control operations, and petrochemical and other process control operations. Anyone who works in, oversees, interacts with, or benefits from a high-consequence work system, such as those just listed, can benefit. When new technology is introduced into high-consequence work systems, high costs are at stake. In some cases, national security is at stake. In many cases, human lives and well-being are at stake. We need tools and methodologies for effectively guiding the entry of new technology into high-consequence complex systems.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The TRUSTS Framework will provide a more productive system-centered approach to the transition process. It will help transition teams to, instead of blaming users or management, focus on the underlying and fundamental sources of tension involving perceived and possible threats to the host system's resilience mechanisms. It will help them to preserve the host system's ability to respond adaptively to high-stakes, time-pressured, and challenging situations throughout and following the transition process.

An anticipated outcome will be that many more technologies will be successfully transitioned every year, hundreds of millions of dollars invested in technology development and upgrades will be viewed as effective versus wasted, envisioned system improvements will be realized, and system modernization programs will experience a smoother process and higher rate of success. Another outcome will be the avoidance of high-risk transition processes and associated stresses on the receiving system, increased rate of near-miss incidents and mishaps, and compromised safety of humans who work in or interact with the system.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Because the TRUSTS Framework is derived from cross-disciplinary theory and principles of the complex-systems science and resilience engineering literatures, it has great potential for generalizability to high-consequence sociotechnical systems across scales and sectors. We have begun investigating the application of the TRUSTS framework to technology transitions in airspace operations, healthcare, and military operations. For the future, we are interested in using TRUSTS to guide the high-level strategic planning and assessment of sociotechnical system redesigns, e.g., of city or national infrastructure for responding to disasters or crises, national and international supply chains, and military strategy that is adaptive in a new international environment. It additionally has potential to be used to evaluate system security and preserve network resilience as new applications are added.

ABSTRACT

The rate at which new information technology fails to successfully transition is disconcertingly high. Many technologies only manage to transition through a forced and difficult process, and often produce system disruptions that pose serious threats to high-consequence systems (e.g., Finkelstein & Dowell, 1996; Goldstein, 2005; Patterson, Militello, Su, & Sarkar, 2016; Sherwood, Neville, McLean, Walwanis, & Bolton, 2020; Trist & Bamforth, 1951; White, Wastell, Broadhurst, & Hall, 2010; Wears, Cook, & Perry, 2006). We propose a systems-centered approach to technology transition and system modernization grounded in complex-systems science and resilience engineering. In an exploratory research effort, we developed an initial version of the Transform with Resilience during Upgrades to Socio-Technical Systems (TRUSTS) Framework. This framework, derived from a literature review, specifies general characteristics of complex systems that allow them to behave with resilience in complex, high-stakes operating environments. We have begun translating the framework into tools for guiding technology transition and system modernization. The toolset will help technology-development and transition teams evaluate and engineer into their designs the system-resilience impacts of new technologies starting with the concept development stage and through deployment. Our primary goal is to use the TRUSTS Framework and tools to significantly improve the success rate of technology transition and count the transition as a success not just because it is being used but because it has transitioned without compromising the receiving system's ability to perform and adapt in high-demand conditions. Complementary efforts will focus on applying and adapting the TRUSTS Framework and toolset to benefit transition and modernization challenges faced in multiple high-consequence domains, including air transportation, healthcare, and national security.

REFERENCES

Finkelstein, A., & Dowell, J. (1996, March). A comedy of errors: the London Ambulance Service case study. In *Proceedings of the 8th International Workshop on Software Specification and Design* (pp. 2-4). IEEE.

Goldstein, H. (2005). Who killed the virtual case file? *IEEE Spectrum*, *42*(9), 24-35.

Patterson, E. S., Militello, L. G., Su, G., & Sarkar, U. (2016). Characterizing a naturalistic decision-making phenomenon: loss of system resilience associated with implementation of new technology. *Journal of cognitive engineering and decision making*, *10*(3), 229-243.

Sherwood, S.M, Neville, K.J., McLean, A.L.M.T., Walwanis, M.M., & Bolton, A. (2020). Integrating new Technology into the complex system of air combat training. In H. Handley & A. Tolk (Eds.), *A framework of Human System Engineering: Applications and case studies.* Washington, D.C.: IEEE.

Trist, E.L. & Bamforth, K. (1951). Defences of a work group in relation to the social structure and of coal-getting: An examination of the psychological situation and some social and psychological consequences of the longwall method and technological content of the work system. *Human Relations, 4*, 3-38.

White, S., Wastell, D., Broadhurst, K., & Hall, C. (2010). When policy o'erleaps itself: The 'tragic tale'of the Integrated Children's System. *Critical Social Policy, 30*(3), 405-429.

Wears, R. L., Cook, R. I., & Perry, S. J. (2006). Automation, interaction, complexity, and failure: A case study. *Reliability Engineering & System Safety*, *91*(12), 1494-1501.

Highlights of Recent Work
WAx on, WAx off: understanding the dynamics of cyber-socio-technical systems

Riccardo Patriarca[a], Andrea Falegnami[a], Francesco Costantino[a], Giulio Di Gravio[a], Antonio De Nicola[b], Maria Luisa Villani[b]
[a] *Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Italy*
[b] *Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy*

- **The WAx (*Work-As-x*) is a conceptual framework that aims to provide a deeper understanding of cyber-socio-technical systems (CSTSs) integrating resilience engineering and knowledge management.**
- **The WAx is relevant for scholars who deal with complex systems from a safety and human factors perspective, but it remains valid for larger resilience-oriented research. It is also expected to be of interest for safety and security practitioners, sociologists, psychologists, and ergonomics and human factors specialists.**

The WAx is a conceptual framework that addresses the needs for understanding, modelling, and analyzing cyber-socio-technical systems, seen as the archetypal maximum expression of complexity achievable by collective systems formed by humans and technological artifacts (Patriarca *et al.*, 2021). The scope of a socio-technical system is expanded by the cyber elements, necessitating their reconsideration and re-evaluation. The introduction of a kind of artifacts and agents (i.e., cyber-agents) objectively different in capability and rapidity of response, raises terrifically the number of possible non-linear interactions within modern systems. To this extent, CSTSs sometimes include open AI issues (e.g., explainability, interpretability, and trust in technological artifacts) and exacerbate features related to adaptation, evolution, and, in general, unpredictability of complex adaptive systems. The understanding, analysis and, ultimately, the engineering of CSTSs, fundamental in the immediate future in terms of safety and security, passes through the definition of an appropriate systemic dimension of investigation. Risks of social manipulation, new types of warfare, or changes in power structures potentially attributable to cyber-artifacts come to mind in this sense. Relying on these observations, the WAx is structured to be impractical for reductionism and simplistic approaches that are patently unsuitable for CSTS risk and safety management (Hollnagel, Woods and Leveson, 2012). Moreover, it aims to adhere to the recently published manifesto for a reality-based safety science (Rae *et al.*, 2020).

Following the methodology proposed by Jabareen (2009) the WAx framework has been developed by a research group formed by scholars in industrial engineering and resilience engineering, and scholars in information science and knowledge management. The research team attempted to connect both domains while keeping a unique systemic view. The WAx is grounded in contributions from several different disciplines (natural and behavioral sciences, resilience engineering, cybernetics, knowledge management theory). In its very essence it is based on the integration of the many varieties of human work as embraced by resilience engineering theory (Moppett and Shorrock, 2018) – though still under-explored in relation to CSTS – with the knowledge management SECI model proposed by Nonaka (Nonaka, Toyama and Konno, 2000).

Succinctly, the framework is articulated in a three-level knowledge structure (CSTS world, CSTS primary knowledge, and CSTS analysis knowledge), along which foundational knowledge conversion activities (*socialization, introspection, externalization, combination, internalisation, conceptualisation, and reification*) take place. The foundational activities create or convert knowledge among six different types of knowledge entities (*Work-As-Done, Work-As-Imagined, Work-As-Prescribed, Work-As-Normative, Work-As-Disclosed, Work-As-Observed*) that are distributed between tacit and explicit types of knowledge. The framework considers an additional dimension of inquiry relative to agency, which is articulated into three generic core agents: the *blunt-end operator*, the *sharp-end operator*, and the *analyst*. Both operators' knowledge belongs to the CSTS primary knowledge (either sharp- or blunt-end), while the analyst's knowledge lies in the CSTS analysis knowledge level. The remaining level, the CSTS world, is where WAD takes place.

The three generic core agents may be, from case to case, organizations, teams, people, artifacts. The WAx framework mimics the recursively fractal nature of a complex adaptive system, for which the component subsystems are themselves adaptive systems, in a system-of-systems perspective. Moreover, the framework encompasses an additional meta-activity relevant for knowledge analysis conversion dynamics, which is *influence*: the effect played by an agent when eliciting knowledge on the foundational knowledge conversion activities. Finally, the framework here presented envisions a set of knowledge conversion drivers: part of them deliberative (so-called ETTO drivers, including *process ETTO transformation, modeling ETTO transformation,* and *communication ETTO transformation*) and a remaining driver *accidental* which represents the inescapable information loss inherent in every physical process.

The WAx framework provides an enabling metaphorical conceptual framework for understanding how a CSTS operates, including the perspective of the observer. It is not an off-the-shelf tool, but rather a conceptual instrument to facilitate critical thinking and raising critical questions by investigators, analysts, and resilience scholars. Various professionals involved in business and administrative processes, such as organizational management or human resources experts could find advantage in the WAx as well, to the larger extent of understanding complex system's variability and performance. Similarly, actors involved in industry (both traditional and 4.0), or service industries such as healthcare (at both management and operational levels) can follow the WAx logic to support knowledge elicitation and knowledge transfer. Scientists involved in complex adaptive systems, systems-of-systems, socio-technical systems, infrastructure as well as experts in human factors, cybernetics, knowledge management, psychologists and sociologists, might find the Wax framework relevant. Going up in abstraction, the framework mimics research questions as they could be raised by scholars concerned with law, epistemology, and philosophy of engineering.

Given the fact that much of the information distributed in CSTSs is embodied in operators or embedded in artifacts - hence tacit, the WAx framework main use is as a mapping device to unveil the knowledge conversion routes hidden within the system analyzed. As an example, one can consider the WAx initial usage as a conceptual support aid for the creation of a software environment meant to facilitate data collection for the Functional Resonance Analysis Method (FRAM). A typical use of the FRAM is comparing the so-called Work-As-Imagined (WAI) and Work-As-Done (WAD) instantiations. In practice, WAI and WAD are both inaccessible and, usually, assumed to be WA-Normative/WA-Prescribed and WA-Observed, respectively, as reasonable proximate measures. While the former are relatively easily accessible, the latter require substantial human and time resources. The idea behind the WAx usage is to build an elicitation tool powered by computer ontologies for crowd-based data collection among frontline operators. The WAx has the potential to describe the influence effects involved in the knowledge transfer process among the various agents of a CSTS. The framework facilitates the formalization of the various FRAM steps and the construction of a representative model for the formalization of the necessary requirements for the software to be implemented. Elsewhere, the WAx can be employed for systemic understanding of phenomena related to systems knowledge management.

In a broader sense, the framework remains a conceptual support for reflection, useful for encouraging discussion and insight into the different system's performance. As such, it casts a light to remove biases on the dynamics involved in the different varieties of work, whether performed by humans or cybernetic agents.

ACKNOWLEDGEMENTS

This work is supported by the H(CS)2I project, which is partially funded by the Italian National Institute for Insurance against Accidents at Work (INAIL) and Institution of Occupational Safety and Health (IOSH) under the 2018 SAF€RA EU funding scheme - H(CS)2I funded project.

REFERENCES

Hollnagel, E., Woods, D. D. and Leveson, N. (2012) *Resilience engineering: Concepts and precepts*, *Resilience Engineering: Concepts and Precepts*. Ashgate Publishing Ltd. doi: 10.1136/qshc.2006.018390.

Jabareen, Y. (2009) 'Building a Conceptual Framework: Philosophy, Definitions, and Procedure', *International Journal of Qualitative Methods*. SAGE Publications, 8(4), pp. 49–62. doi: 10.1177/160940690900800406.

Moppett, I. K. and Shorrock, S. T. (2018) 'Working out wrong-side blocks', *Anaesthesia*. Blackwell Publishing Ltd, 73(4), pp. 407–420. doi: 10.1111/anae.14165.

Nonaka, I., Toyama, R. and Konno, N. (2000) 'SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation', *Long Range Planning*. Pergamon, 33(1), pp. 5–34. doi: 10.1016/S0024-6301(99)00115-6.

Patriarca, R. *et al.* (2021) 'WAx: An integrated conceptual framework for the analysis of cyber-socio-technical systems', *Safety Science*, 136. doi: 10.1016/j.ssci.2020.105142.

Rae, A. *et al.* (2020) 'A manifesto for Reality-based Safety Science', *Safety Science*, 126. doi: 10.1016/j.ssci.2020.104654.

Highlights of Recent Work

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Mohammadreza Jalaeian (Reza)[a], Peter Fadde[b]
[a]*ShadowBox Training LLC/MacroCognition LLC*
[b]*Southern Illinois University Carbondale*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

- This line of research involves the use of a novel approach that focuses on investigating design and development of systematic design of instruction–type training programs that can accelerate expertise of law enforcement officers.
- The methodology can be extended into other domains such as medicine, military, aviation, and business and create training programs that are cheap, fast, easy, and flexible for professionals.

2. Summarize New results

Law enforcement officers often must make rapid decisions and actions to protect themselves and control a situation or a subject. While different agencies teach different techniques and philosophies for use-of-force and arrest-and-control, they are all enhanced by accurate and fast recognition of potentially attacking motions made by subjects. Whether subjects are civilians being questioned or suspects being detained, officers must respond with an appropriate scale and type of action. Slow or inappropriate actions can have highly negative results for subjects, officers, and their agencies. Unfortunately, training of rapid decision-action performance requires a high volume of practice that is difficult to acquire in training contexts that typically focus on “what you do” more than “what you see.”

One possible solution is to target the Recognition component for concentrated practice using the Video-Occlusion method developed in sports science. Video-occlusion was validated by reliably differentiating expert performers from less expert performers. It was then developed into video-occlusion training applications. For decades researchers have observed that video-occlusion might contribute to training rapid decision-action performance in military and law enforcement contexts.

This study compared the performance of experienced law enforcement officers and less experienced officers on a video-occlusion test of Attack Recognition. Participants in the study viewed 48 video examples of attacking and non-attacking movements by subjects. Video clips were edited to black (occluded) at various points in subjects’ movements, including occlusion just prior to the motion commencing. Participants identified the type of attacking (overhand punch, underhand punch, kick, or knife draw) or non-attacking motion.

Findings validate the video-occlusion method and recommended specific occlusion points for recognition-only video-occlusion training of attack recognition that can be delivered on tablets and cell phones as a way to accelerate expertise.

3. Who would benefit from knowing about your work

This work might would benefit professional in various domain including law enforcement, military, health care, and many more. Those who need to make decision fast and under constraints of time and stress mighy benefit the most.

4. The impacts or implications of the results or how they can be utilized

This is the first study that employs video-occlusion method to test law enforcement officers’ perceptual–cognitive skills, specifically the sub-skill of attack recognition. As such, it has immediate value as a screening test of whether officers are adequately prepared for assignments that involve interactions with civilian subjects that may require use of defensive tactics to control situations and protect citizens as well as themselves. Beyond use as a pre-assignment screening test, this study provides a strong foundation for recognition-specific perceptual training of reactive skills in law enforcement and many other domains.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The present of focus research and training on a specific perceptual–cognitive skill, such as attack recognition, is very fine-grained. As recommended in a recent review of perceptual–cognitive expertise in sport (Williams, 2020), validation-level studies should move to multidimensional environments that facilitate accelerated development of perceptual–cognitive expertise. Combining research in instructional design, expertise studies, and law enforcement areas, multiple dimensions can include biases (e.g., race), psychological and physiological factors (e.g., stress response), gaze direction (e.g., recognition of Non-verbal Cues of Interpersonal Violence (NCIV). Ultimately, multidimensional Part-task approaches may offer many benefits of whole-task training while maintaining development and learning efficiencies associated with systematic design of instruction approaches.

Running out of Resilience: coping with the COVID-19 pandemic in Brazilian intensive care units

Tarcisio A. Saurin[a], Priscila Wachs[b], Wagner P. Bueno[a], Carolina Zani[a], Ricardo Kuchenbecker[c], and Márcio M. Boniatti[d]
[a] *Universidade Federal do Rio Grande do Sul, Industrial Engineering Department, Brazil*
[b] *Instituto Federal de Educação do Rio Grande do Sul, Brazil*
[c] *Universidade Federal do Rio Grande do Sul, Faculty of Medicine, Brazil*
[d] *Hospital de Clínicas de Porto Alegre, Brazil*

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

Intensive care units (ICUs) have been the most visible facet of healthcare services affected by the COVID-19 pandemic, in which this finite resource has been the ultimate battleground for struggling people and systems. Despite media reports and personal accounts by practitioners, there is a need for systematic investigations of how ICUs have coped with the pandemic from a non-clinical, organizational perspective.

2. Summarize New results

This paper presents an exploratory investigation of how ICUs in Brazil, one of the countries most affected by the pandemic, have coped with it. A resilience engineering perspective is adopted, using a mixed-method research design. The quantitative stage comprised an exploratory survey for the assessment of the uptake of five work system guidelines that are logically related to resilience and address the micro and meso levels. The guidelines and their corresponding number of survey questions are: provide slack resources (6); give visibility to processes and outcomes (4); support diversity of perspectives when making decisions (4); monitor and understand the gap between work-as-imagined and work-as-done (5); and monitor unintended consequences of improvements and changes (4). These guidelines are described as statements in the survey, accompanied by ICU examples. There is a sliding bar with two endpoints: fully disagree (corresponding to zero) and fully agree (100). For example, one of the statements related to the guideline on slack is as follows: *there are extra or standby human resources that can be quickly deployed, and these are available in sufficient quantity to cope with unforeseen events*. There were 33 valid responses, all from different ICUs, from 13 out of the 27 Brazilian states.

All survey respondents were invited to be interviewed and seven agreed with that. Interview questions stemmed from the survey statements and they shed light on the rationale for the survey responses. The qualitative stage followed with interviews with two public health officials and 20 hours of non-participant observations of the meetings of a municipal COVID-19 crisis management committee in a capital city in Southern Brazil – these sources of data contributed to the assessment at the macro level. Joint data analysis from the quantitative and qualitative stages supported the identification of five lessons learned, each corresponding to one of the aforementioned guidelines. Data collection took place from September 2020 to January 2021, while the pandemic was raging in Brazil, although not as dramatically as in March 2021.

Results indicated the guideline on slack resources as the most relevant as it addressed the key pandemic issue of matching capacity to demand. In this respect, the provision of sufficient and experienced staff was found to be more challenging than the provision of facilities and supplies. In addition, results indicated that running out of resilience was part of everyday work in the ICUs, as evidenced by: (i) the lack of effective treatments, which means that despite all efforts the mortality rates were very high; (ii) exhausted healthcare professionals who could not cope with the stressful working conditions – there were many reports of professionals who resigned; and (iii) chronic and acute mismatches between capacity and demand, which implied that many patients received sub-standard care. Furthermore, it was clear that ICU resilience was dependent on the resilience of society at local, national, and international levels – in fact, ICU resilience was mostly a necessary reaction to variability from the external environment. This suggests that an integrated approach to resilient healthcare is necessary, encompassing the micro, meso, and macro levels, besides accounting for the broader societal resilience.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

The target audience of this paper is formed by researchers and practitioners interested in resilient healthcare and human factors. They will obtain insight into how ICUs in Brazil coped with an unprecedented disruption (i.e., COVID-19 pandemic), from an organizational, non-clinical viewpoint. The lessons learned are likely to be applicable to other countries and sectors as the pandemic has had a worldwide impact.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The adopted research method can be replicated, partly or totally, in other countries for the investigation of the pandemic implications. For example, the survey questionnaire can be applied to a larger national and international sample of ICUs, allowing for the identification of benchmarks. Also, the lessons learned can be useful for the improvement of existing resilience engineering tools. These tools could place more emphasis on issues that proved to be a key for coping with the pandemic - e.g.; provision of a wide mix of slack resources, and putting a stronger emphasis on the anticipation of threats and societal resilience.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

The COVID-19 pandemic is still unfolding and its impacts will be felt for years. It is a highly complex phenomena that has implications not only for the resilience of health services but for society at large. This work is a step towards the understanding of those implications from the lens of resilience engineering. Existing practices and assumptions of resilience engineering should be revisited in light of the practical and theoretical lessons learned from the pandemic. In particular, the collapse of health services and businesses in Brazil and in several other countries raises questions on how to expand the limits of resilient performance during a major and prolonged international crisis.

The Workload Implications of Resilient Healthcare: the Role of Social Interactions

Stela X. Terra[a], Tarcisio A. Saurin[a], Flávio S. Fogliatto[a] and Ana Maria M. Magalhães[b]
[a] *Universidade Federal do Rio Grande do Sul, Industrial Engineering Department, Brazil*
[b] *Universidade Federal do Rio Grande do Sul, Nursing Department, Brazil*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD PARTICIPATE AND USE YOUR WORK

The extra effort of clinicians´ to provide care to their patients, which is a manifestation of resilient performance, usually goes unnoticed due to successful outcomes that occur most of the time. Therefore, clinicians' self-sacrifice and overreliance on resilience becomes "normal work". This study investigates the workload implications of resilient healthcare.

2. SUMMARIZE NEW RESULTS

The study setting was the radiology department of a major teaching hospital. The centrality of each clinician in four ability-based social networks (i.e., the resilience abilities of monitoring, responding, anticipating, and learning) was considered as a proxy of the contribution of each actor to the overall system resilience. As such, a resilience score was calculated for each actor in each ability-based network, combining five indicators theoretically connected to the actor's resilience: in-degree, closeness, and betwenness, which are derived from social network analysis, and availability and reliability, which are non-network attributes assessed through Likert-style questions. In turn, the individual workload was assessed based on the NASA-TLX questionnaire, which produces indicators related to six dimensions of workload (in addition to an overall score), namely: mental, physical, temporal, performance, effort, and frustration. Both questionnaires – social network analysis and NASA-TLX – were answered by 155 out of the 230 clinical and non-clinical staff of the radiology department. Follow-up interviews with 10 respondents were carried out for understanding the influence of contextual factors on the actors´ workload and their resilience score based on network centrality. Preliminary data analysis indicates that the overall workload of the actors with the higher and lower resilience scores is similar, although there are significant differences in the workload dimensions (e.g., actors with higher resilience scores are subject to higher temporal and mental demand in comparison to those with lower resilience scores). As another finding, mental demand was the only workload dimension significantly correlated (95% level of confidence) to the resilience score. Furthermore, actors that take the initiative to help their coworkers without being requested had a higher overall workload than those actors that are more passive – the frequency at which actors take the initiative to help others was assessed by a specific question in the social network analysis survey. Overall, our findings suggested that social interactions related to the four resilience abilities mattered to the workload of the professionals surveyed.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

The target audience of this paper is formed by resilience engineering and human factors researchers and practitioners. They will obtain insight into the role of social interactions in the workload of healthcare professionals, which in turn can set a basis for work system redesign initiatives.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS OR HOW THEY CAN BE UTILIZED

The adopted research method is expected to be replicable to other settings. It might be a new approach for the analysis of the human costs of resilient performance. Workload is proposed as a proxy of that human cost.

5. THE WIDER SCOPE OR RELEVANCE OF THE WORK

Although resilient performance is necessary in complex systems, earlier studies suggest that it is commonly overused and not properly supported by organizational design, relying too much on individual performance. Despite this, to the best of our knowledge, no previous study was concerned with the modelling and quantification of the resilience implications on workload. Therefore, the present work potentially opens a new stream of resilience engineering research, which can explore the human costs of resilience from a number of perspectives and methods in addition to those described in this paper. For example, the relationship between resilience and burnout can be worth investigating as well as the relationship between resilience at the individual and psychological level with the resilience score proposed in this study.

Interactive sessions

Interactive and Integrative Sessions

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Kees van der Blom[a], Jurriaan Cals, John van Schie, Christiaan Poll and Anne van Galen
[a] *Blom Safety*

1. Two sentence teaser that succinctly communicates why someone should attend and participate in this session

- Serious gaming ……..honestly, how serious can gaming be?
- And how do we proceed with the idea of resilience?

2. Summarize your plan for the session

Over the past few years, we developed suitable theoretical material for this purpose and also, great efforts have been made to bring resilience to the shop floor using the concept of a game.

Being an (ad hoc) working group we are committed to resilience (and everything that goes with it). That is why we would like to use the existing theory and build on the concept of a game; we have the ambition to make resilience easily accessible for management and the shop floor. This means that we want to work on a game concept that people want to play and where, in competition with others, they opt for resilient choices and at the same time learn by having fun.

In short, we think there is enough reason to set off (again) with resilience and there is enough demand. We believe that serious gaming can be an important tool to use for working with resilience theories on the shop floor. We are still in the development stage of our serious game and in the process we noticed that we were confronted with several interesting choices. We would like to share these experiences with you in a workshop specifically developed for the REA conference. The workshop serves a dual purpose: apart from discovering which problems we may encounter, this may also provide us input for the concept choices we must and intend to make with the game.

The starting points for setting up this experiment are:

- to learn how to build a good game based on a serious game framework;
- to experience how a serious gaming concept can help in developing operational resilience and team resilience on the shop floor;
- subsequently to experience how gaming can help to make a start in using resilience features on the shop floor; and
- finally, with the help of serious gaming, how to work out the resilience concept together.

And as said, it must definitely be fun and interactive too if you want to work with serious gaming in this way.

3. The wider scope or relevance of this session

For us being an ad hoc working group, this is a serious effort to jointly develop a serious game in an interactive and fun way and which game we, working group, may use as basis to proceed on.

It would be good if in the end we could work together and build a serious game with which we can roll out the principles of team resilience and operational resilience on the shop floor and thus give safety a positive impetus.

Interactive and Integrative Sessions

From critical infrastructure to society as a whole: Can Naturalistic Decision Making and Resilience Engineering scale up?

Session Chair, CoChair: Matthieu BRANLAT[a], Johan BERGSTRÖM[b]
Participants: Jan VERLIN[c], Martina RAGOSTA[a], Francesca de'Donato[d], others TBD
[a]*SINTEF Digital, Trondheim, Norway*
[b]*Lund University School of Aviation, Lund, Sweden*
[c]*Chaire Géopolitique du risque, Ecole Normale Supérieure, Paris, France*
[d]*Department of Epidemiology Lazio Regional Health Service, Rome, Italy*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD ATTEND AND PARTICIPATE IN THIS SESSION

- The current pandemic emphasizes the role the population (individuals, informal groups) plays in the management of crisis situations and the necessity to consider society as a whole for building resilience.
- The panel brings researchers and practitioners from multiple disciplines to explore how concepts and methods of Resilience Engineering and Naturalistic Decision Making can be mobilized to better understand how informal social actors make decisions during crisis that lead to resilient societies.

2. SUMMARIZE YOUR PLAN FOR THE SESSION

In light of events such as the Fukushima accident or the current pandemics, disaster management has been shifting in recent years from a "critical infrastructure" to a "societal" focus. In RE and NDM, we are used to deal with the former, the more professional or formal side of management of disruptive events, through our concepts and methods: we investigate, learn from and try to better support expert behaviour or resilient organisations/HROs. The potential and limits of applying our analytical tools to informal social actors remains less studied.

The proposed panel aims to address questions such as the following:

- Are we, as fields of research, equipped to look at the larger societal context and dynamics?
- How can we account for non-professional individuals and groups actions?
- Should we, as a field, be involved in this larger scale, or should we only ensure we can interact with other fields who look more broadly? In the latter case, what can we bring to such discussions?
- What additional competences do we need to bring in or what fields do we need to be inspired by in order to be able to ask better questions on societal scales?

The discussion will be chaired by two researchers who have been involved in RE and safety research for years, and will involve speakers from different background:

- researchers in political science, sociology, community resilience, and resilience engineering
- practitioner from organisations of the health sector

A few other speakers will be identified prior to the event, especially with a research or practice focus on NDM, societal resilience and civic participation.

We propose to implement a dialogue-oriented panel, in which the discussion is initiated from a pre-identified set of speakers presenting shortly different perspectives on a few prepared questions. Attendants will be invited to join, similarly to a fishbowl session.

3. Who would benefit from participating in this session

RE and NDM researchers curious about the topic of disaster management, and about discussing the scope of their fields of research.

Researchers and practitioners curious about what concepts and methods RE and NDM can bring to the understanding of societal resilience.

4. How the outcome of this session might be utilized (i.e., will this session result in a paper/book, shape future research, promulgate a viewpoint/perspective/method, etc.)

The session aims to shape future research and inform the audience of societal perspectives around the notions of risk, participation and resilience.

5. The wider scope or relevance of this session

The notion of resilience has gained tremendous popularity in recent years, as reflected by policy documents such as the UN sustainability goals or the SENDAI framework, by research funding calls, or by the general language used around the disruptive events our societies face. RE and NDM are relevant fields in this context and have an opportunity to be provide concepts, insights and methods to the large questions around societal resilience.

Please limit your submission to two pages or less

Interactive and Integrative Sessions--Panel

An Intelligent Assistant You Won't Want to Fire

Session Chair and presenter: *Cindy DOMINGUEZ*, the *MITRE Corporation*
Co-chair and discussant: Robert HOFFMAN, Institute for Human and Machine Cognition
Presenters: John MERRIHEW, Veloxiti; Patty McDermott, the MITRE Corporation;
Kerstin Haring, University of Denver

1. TWO SENTENCE TEASER:

- People query intelligent assistants on personal devices to help with ever-expanding tasks and decisions throughout their day, yet it is difficult to imagine these assistants being reliable or effective enough to support decision makers in NDM-like settings.
- What is the current state of research applying this technology to classic NDM situations, and what are the potential payoffs and problematic pitfalls of a JARVIS[1] for high-stakes decision makers?

2. SUMMARIZE YOUR PLAN FOR THE SESSION

a. Format: Panel session, where the session chair and panelists will each provide a 10 minute overview of their research and/or perspective on this topic, with time for discussant and facilitated audience discussion after each panelist's talk as well as an additional half hour for open discussion after presentations. We will have four panel presenters and one discussant.

b. Panelists will describe their research, applications, and summative recommendations on the interaction between real world, sharp-end decision making and what key near- and far-term research should be undertaken to realize an "intelligent" assistant within the sphere of narrow AI and automation. Discussion of partnership paradigms and sharable design patterns based on applied research will help the audience to translate the ideas into their own domains of research.

Speaker background:

- **Dr. Cindy Dominguez** is a Principal Cognitive Engineer and Capability Lead for Human Machine Teaming (HMT) at the MITRE Corporation. She is working to accelerate methods, processes, and outreach for helping government and industry develop and adopt intelligent systems, and to apply NDM methods on a daily basis in project work. She served as a US Air Force officer for 20 years. Her PhD, earned in 1997, is from Wright State University in Human Factors. She has worked on applications to assist soldiers to easily direct incoming unmanned supply helicopters, and on recent design patterns for a command and control IA.
- **Dr. Robert Hoffman** is a recognized world leader in cognitive systems engineering, NDM, and Human-Centered Computing with expertise in the psychology of expertise, cognitive task analysis methodology, and human-centering issues for intelligent systems technology. He is a Senior Member of the Association for the Advancement of Artificial Intelligence and numerous other professional organizations, and has authored several books and recent publications directly relevant to this panel's topic. His current work focuses on "Explainable AI."
- **John Merrihew** has served as the Principal Investigator on a dozen different Artificial Intelligence projects in the DoD in the domains of data analytics, operations/ intelligence fusion, UAV automatous operations and human terrain. Mr. Merrihew is a retired Lieutenant Colonel Army Aviator who served 24 years in the United States Army in numerous positions. He is a 4 time combat veteran of Operation Iraqi Freedom with 113 combat sorties and over 1300 combat hours serving as Standardization Instructor in the 116th Air Control Wing. He has had command and staff positions at multiple levels in the US Army with two tours on the DMZ in South Korea. He served in Infantry, Field Artillery, and Aviation (Cavalry, Attack and

[1] Marvel's Iron Man series employed a fictional AI called JARVIS (Just a rather very intelligent system).

Lift) units. John has extensive practical combat experience in intelligence fusion for warfighters at the tactical edge and is qualified in Joint Operational Planning.

- **Patty McDermott** is a Principal Cognitive Engineer at the MITRE Corporation with over 20 years of experience in naturalistic decision making. She regularly conducts cognitive task analyses to understand users and to improve systems, training, or models. She specializes in decision support, human-machine teaming, visualization, and the evaluation of systems to assess how well technology supports decision making and situation awareness. She has a MS in Human Factors Engineering from Wright State University. She is currently developing an intelligent assistant application to support government acquisition managers.
- **Kerstin S. Haring** is an Assistant Professor of Computer Science at the University of Denver, leading the "Humane Robot Technology" lab there. Before this, she was a researcher in Human-Machine-Teaming at the U.S. Air Force Academy. She completed her PhD in Human Robot Interaction at the University of Tokyo in Japan and her Master in Computer Science and Cognitive Science at the University of Freiburg in Germany. Her work focusses on evaluating and enhancing collaborative robot capabilities and increase the interaction effectiveness of social robots, examining trust, collaboration, ethical implications, and teaming with robots and other autonomous systems.

c. Questions to be explored:

- What does adaptive behavior look like for an intelligent assistant?
- What are system-level affordances as well as limitations and pitfalls of the concept of an intelligent assistant?
- How can system designers build calibrated trust into their assistant capabilities?
- What is the current state of the art vs the potential state of the art based on the rapidly evolving state of AI technology?

d. Outcomes. Attendees will have an understanding about current thinking and approaches in the development of assistants towards future resilient, adaptive performance in high-stakes settings.

WHO WOULD BENEFIT FROM PARTICIPATING IN THIS SESSION

- NDM/REA practitioners who can share, or would like to learn about, state of the art challenges and successes in Intelligent Assistant design and development.

3. HOW THE OUTCOME OF THIS SESSION MIGHT BE UTILIZED

This session is intended to generate collaboration across panelists and attendees towards a book chapter summarizing the current state of the art and outstanding questions needing to be addressed towards progress in Assistant/Associate development, as well as case studies and success/failure stories.

4. THE WIDER SCOPE OR RELEVANCE OF THIS SESSION

Researchers have been working on 'associate' or 'assistant' technology for decades. Current software capabilities are mature enough to provide assistance in various context-specific ways. In the organizer's recent work envisioning near-term software modernization, the idea of a JARVIS-lite visual software based entity has been immensely appealing to operational experts and technologists. The NDM/REA community should collaborate on ways to get ahead of the next generation of assistants with guidance and patterns for command and control, healthcare, and other domains. This session is meant to start the community discussion and collaboration towards improving future assistant technology.

Interactive and Integrative Sessions
Surprises and adaptations in software development projects

Michael W. SMITH[a], Jordan M. HAGGIT[a], Marisa R.G. BIGELOW[a]

[a]Mile Two, LLC, Dayton, OH, USA

1. Two sentence teaser that succinctly communicates why someone should attend and participate in this session

- There is understandably much emphasis on ways to prevent surprises in software project management, but the corollary - how to adapt to surprises that do occur - has been neglected.
- This interactive session will explore examples of surprises in software projects (our own and those contributed by conference participants), focusing on resilience and associated macrocognitive functions.

2. Summarize your plan for the session

The design and development of software systems is a complex process. Even a relatively small project may involve distributed teams of several different types of professions, dozens of different tools, and hundreds of different documents. Prescriptive system development processes, of which there are many, are focused on how to proactively reduce potential variation. By contrast, there has been little work exploring how to adapt to unanticipated variation - how to cope with surprises - in software project management.

Surprises in the systems development process can impact the cost of development, the scope or quality of the product, the time of delivery, and the relationships between stakeholders. They can have positive or negative impacts, opening new opportunities or hindering progress forward.

Responding to a surprise constitutes a test of the ability of the project team and parent organization to recognize and adapt. The sensitivity to meaningful variation and potential risks, the flexibility of work processes and resource allocation, and the expertise of the responsible practitioners all play important roles in the adaptation process.

In this session we will begin by discussing patterns in problem detection (Klein), anomaly response (Woods & Hollnagel), and organizational surprises (Cunha, Clegg & Kamoche). We will share some examples from our own work at Mile Two and analyze them. The aim of the analysis will be to identify macrocognitive functions related to coping with surprises in this context, and how project or organizational resilience is evidenced in these examples.

To make this an interactive session, at the start of the conference we will ask participants to share electronically their own accounts of surprises in systems development projects. We will include these in our analyses - ahead of time for earlier contributions, and in real time for late or concurrent contributions. We will use an online whiteboard or similar tool to actively show the real-time analysis.

3. Who would benefit from participating in this session

People working in systems development, those interested in problem detection, replanning, and/or organizational resilience.

4. How the outcome of this session might be utilized (i.e., will this session result in a paper/book, shape future research, promulgate a viewpoint/perspective/method, etc.)

The material generated as part of the activity will be the basis of a paper.

5. The wider scope or relevance of this session

Problem detection and replanning (detection of and response to anomalies) constitute key macrocognitive functions and are essential for resilience.

Interactive and Integrative Sessions

Workshop: Designing for Human-Machine Teaming – Beyond the Buzz

Patricia MCDERMOTT, Cindy DOMINGUEZ[a]
[a]*The MITRE Corporation*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD ATTEND AND PARTICIPATE IN THIS SESSION

- Do you want to develop autonomous or AI systems that partner with humans by expanding user thinking, providing backup when needed, and being able to explain what and how the technology is doing?
- In this hands-on interactive workshop, facilitators will cut through the jargon and hype to provide actionable guidance and methods for developing autonomy, automation, and AI that partner effectively with humans.

2. SUMMARIZE YOUR PLAN FOR THE SESSION

In this workshop, we will start by unpacking the Human-Machine Teaming (HMT) Framework, which provides a language for discussing HMT and the eight key themes for designing and developing HMT. The instructors will provide an overview of a research-based HMT Framework that addresses the need for appropriate observability, cognitive assistance, and coordination.

Participants will practice a set of methods for eliciting HMT needs and designing interfaces and interactions with emerging autonomy, AI, and automation-based systems. They will learn how to develop and apply HMT design guidance to a specific system or application. Techniques will be brought to life with exercises to practice gathering HMT information from experts, analysing interview data, applying findings to design, and conducting HMT evaluations. Participants will use the HMT Knowledge Audit to elicit HMT considerations from an expert and then use the data collected to develop user stories that illustrate HMT needs in context. Depending on the time allotted to the workshop, participants will use paper and pencil to design interfaces between humans and automation/autonomy/AI and plan evaluation techniques that can assess how well the HMT needs are supported.

The workshop will be highly interactive with lively discussions and exercises to practice HMT methods in context. Discussion topics include:

- Perspectives on trust – do we need to get the user to trust the system?
- What are your experiences with Observability in the apps/tools you use?
- Should AI be able to completely explain itself?
- Discussion of the myth that we need to replace the overworked unreliable human. How have you successfully countered this perspective?

In addition to the knowledge elicitation practice, participants will dissect the 737 Max incident to identify ways in which HMT was not supported and suggest ways in which HMT could have been supported. This provides an opportunity for participants to further apply what they have learned.

The workshop is designed to be a full day course but can be modified to be a half-day session if needed. Upon completing the workshop, participants will be able to:

- Understand the wealth of research in HMT and how to use it towards effective design;
- Advocate for AI as a partner vs. a tool;
- Discuss possible roles for humans when teaming with automation/autonomy/AI;
- Apply HMT systems engineering methods to develop automation/autonomy/AI that partners effectively with humans.

Facilitators:

- *Patty McDermott*. Patty is a Principal Cognitive Engineer at the MITRE Corporation with over 20 years of experience in naturalistic decision making. She regularly conducts cognitive task analyses to understand users and to improve systems, training, or models. She specializes in decision support, human-machine teaming, visualization, and the evaluation of systems to assess how well the technology supports decision making and situation awareness. She has a MS in Human Factors Engineering from Wright State University.

- *Cindy Dominguez.* Cindy is a Principle Cognitive Engineer and Human Machine Teaming Capability Lead at the MITRE Corporation. She has over 30 years of experience in which she has researched and applied macrocognitive concepts across a wide range of domains with NDM characteristics. She has worked over her career to systematize the application of cognitive engineering methods and processes to real-world engineering development. She served in the Air Force for 20 years, retiring as a lieutenant colonel, and has been working in Industry for the past 14 years. Her PhD in Human Factors is from Wright State University.

3. WHO WOULD BENEFIT FROM PARTICIPATING IN THIS SESSION

- Researchers, engineers, designers, and project leaders who are designing and evaluating advanced automation, autonomy, or AI

4. HOW THE OUTCOME OF THIS SESSION MIGHT BE UTILIZED (I.E., WILL THIS SESSION RESULT IN A PAPER/BOOK, SHAPE FUTURE RESEARCH, PROMULGATE A VIEWPOINT/PERSPECTIVE/METHOD, ETC.)

The workshop will provide the actionable guidance and tools to apply HMT in the design and evaluation of advanced systems. This will increase the practitioners' ability to advocate for HMT and design to support it. Participants will leave the session with more tools in the repertoire for eliciting, understanding, and addressing specific HMT needs.

5. THE WIDER SCOPE OR RELEVANCE OF THIS SESSION

NDM practitioners will have more tools in their toolset for eliciting, understanding, and addressing HMT needs specific to the system they are designing or evaluating. The ultimate result is automation/autonomy/AI that partners more effectively with humans and thus is adopted, appropriately used, and effective in meeting the mission or goal. NDM practitioners will then be able to feed lessons learned back to the community, increasing the knowledge and research base.

ABSTRACT

Do you want to develop autonomous or AI systems that partner with humans by expanding user thinking, providing backup when needed, and being able to explain what and how the technology is doing? In this hands-on interactive workshop, facilitators will cut through the jargon and hype to provide actionable guidance and methods for developing autonomy, automation, and AI that partner effectively with humans.

Facilitators will unpack the Human-Machine Teaming (HMT) Framework, which provides a language for discussing HMT and the eight themes for designing and developing HMT. Participants will practice a set of methods for eliciting HMT needs and designing interfaces and interactions with emerging autonomy, AI, and automation-based systems. They will learn how to develop and apply HMT design guidance to a specific system or application. The workshop will be highly interactive with lively discussions and exercises to practice HMT methods in context.

Interactive and Integrative Sessions

Facilitated Discussion: Is Trust Overblown?

Patricia MCDERMOTT[a]
[a]*The MITRE Corporation*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD ATTEND AND PARTICIPATE IN THIS SESSION

- There is no audience, by attending you are part of the (large) roundtable so come lend your expertise and opinions!
- In this facilitated session, the group will explore research on trust and calibrated trust; discuss whether the emphasis on 'trusting the system' is overblown; and share lessons learned on evaluating user trust and assessing whether users understand and use the system appropriately.

2. SUMMARIZE YOUR PLAN FOR THE SESSION

- This session embraces full audience participation. Attendees will take the role of participant and not merely observe. Instead of inviting a set of experts to present, this session invites the audience to fully participate as community members with relevant experience and informed opinions. This is an opportunity to hear from a diverse set of voices in the field of NDM.
- Sticky notes (real or virtual) will be used to gather research and opinions. The facilitator is experienced at using a remote collaboration tool, MURAL, to conduct distributed sessions to brainstorm, define problems, and shape solutions. MURAL functions as a digital whiteboard to which participants can add sticky notes, move and group them with similar ideas, draw, and illustrate connections between ideas. MURAL will be utilized if the conference is remote or hybrid to ensure maximum participation. If the conference is fully in person then traditional white board and sticky notes will be used.
- The approximate session length is 90 minutes. This will allow time to introduce the topic, lay the ground rules for participation, and provide a quick MURAL tutorial (if hybrid/remote).
- Patty McDermott will give a short presentation about her experience and challenges in Evaluating Trust in Autonomous/AI systems. This is a challenging area because trust is a buzzword and sponsors and project leads specifically ask for it. There is a critical difference between trust and calibrated trust, which will be explored in the session. Compounding the problem is the fact that many assessment techniques focus on trust at the expense of calibrated trust.
- The bulk of the time will be spent facilitating a discussion about the role of NDM researchers and practitioners in the development and evaluation of trusted and trustworthy systems. Sample discussion points include:
  - What does it mean to 'trust' a system? What are we evaluating?
  - A project lead or sponsor requests that you "increase trust in the system." How do you respond?
  - How do you know that users are trusting the system appropriately? What 'trust' needs are you seeing from sponsors/customers?
  - What is the difference between trust and trustworthiness?
  - What tools, techniques, and metrics have you used?
  - Which are the most diagnostic? Can you provide examples?
  - What gaps remain? As an evaluator, what do additional tools or techniques do you need?
  - What advice would you give someone just out of college who is tasked with evaluating trust?
- A variety of facilitation techniques will be employed to keep the participants engaged and ensure that participants have a chance to have their voice (or pen!) heard. In a variant of Bodystorming, participants will be asked to get out of their seats and based on their opinion/perspective, choose their position on a continuum between two extremes. This is a quick way to get a pulse of the crowd and identify potential speakers with differing opinions. If virtual, this can also be done with sticky notes on a literal continuum line, anchored with defined endpoints. Some questions lend themselves to verbal discussion. Follow up will encourage participants to add on to a comment with a "yes, and" statement or to disagree and state a counterpoint. Other questions would benefit by having a record of the different input and will be captured on the whiteboard. Participants will be given time to individually brainstorm and add their idea (and sticky) to the conversation.
- Outcomes: Multiple NDM perspectives on trust will be elicited. But the primary benefit is the sharing of lessons learned in evaluating trust – how to define it, how to measure it, what gaps exist, and how to reshape teammates perspective and attitude toward trust.

3. WHO WOULD BENEFIT FROM PARTICIPATING IN THIS SESSION

- Researchers and developers working on autonomous or AI systems who have experience designing for [calibrated] trust
- Those who want to learn more about the issues and considerations of *Evaluating Trust and Trustworthiness* from those in the trenches

4. HOW THE OUTCOME OF THIS SESSION MIGHT BE UTILIZED (I.E., WILL THIS SESSION RESULT IN A PAPER/BOOK, SHAPE FUTURE RESEARCH, PROMULGATE A VIEWPOINT/PERSPECTIVE/METHOD, ETC.)

There will be a record of participant input and discussion (in the form of physical or virtual stickies and whiteboard notes). This could form the basis of a practitioner-focused paper on trust assessment. In addition, the identification of gaps could be used to spur research endeavours to fills those gaps.

5. THE WIDER SCOPE OR RELEVANCE OF THIS SESSION

This facilitated session on trust will fulfil two goals. Attendees will walk away with concrete techniques for evaluating trust as well as advice for employing them. They will also have the opportunity to interact and connect with a wide variety of trust researchers that they may not have otherwise met.

ABSTRACT

Trust is a buzzword in autonomous system design, with sponsors and managers often asking for it by name. The waters get murky as practitioners navigate the differences between trust and calibrated trust and the assessment techniques associated with each construct. In this session there is no audience - by attending you are part of a (large) roundtable so come lend your expertise and opinions! Through guided facilitation, the group will explore research on trust and calibrated trust; discuss whether the emphasis on ‘trusting the system’ is overblown; and share lessons learned on evaluating user trust and assessing whether users understand and use the system appropriately. A variety of facilitation techniques will be used to keep participants engaged and ensure that a diverse set of voices is heard. Attendees will walk away with concrete techniques for evaluating trust as well as advice for employing them. They will also have the opportunity to interact and connect with a wide variety of trust researchers that they may not have otherwise met.

Defining an Integrative Framework of Sensemaking and Sustainability for Building Organizational and Community Resilience

Mark CHUBB[a], Charles JENNINGS[b]
[a]*Adjunct Lecturer, Department of Security, Fire & Emergency Management, John Jay College of Criminal Justice, City University of New York, New York, NY, USA*
[b]*Associate Professor and Director of the Christian Regenhard Center for Emergency Response Studies, Department of Security, Fire & Emergency Management, John Jay College of Criminal Justice, City University of New York, New York, NY, USA*

ABSTRACT
Systemic vulnerabilities that cause or contribute to the outcomes of disasters fall under the broad public policy rubric of wicked problems that evade simple solutions. Taming wicked problems requires crisis leadership that addresses challenges and issues through relationships not just resource control and capacity. Conventional approaches to emergency preparedness and planning emphasize a deficit-based approach focused on resource control and capacity. This workshop explores a new paradigm grounded in relationships and informed by public value to align efforts and co-produce public services that mitigate organizational and community risks and expedite response and recovery operations. The proposed framework identifies five virtues associated with successful sustainability initiatives to engage stakeholders' sense of common purpose, shared experience, and creativity. By recognizing individual interests and unique capabilities, the proposed engagement process seeks to foster alignment to leverages creative and collaborative solutions to emergent issues.

KEYWORDS
Civic engagement; co-production; public participation; public value, resilience.

INTRODUCTION

Organizations and communities are embracing the principles and practices of sustainability and resilience in response to wicked problems (Churchman, 1967; Rittel, 1973; Australian Public Services Commission, 2007), such as climate change, social and economic vulnerability, and health and wellness disparities. Leaders often find stakeholders' aspirations and interests conflict even when they are in violent agreement on the need for decisive action to meet such challenges. This workshop proposes a public engagement framework for reconciling stakeholders' conflicting conceptions of the public good to enhance community resilience.

SUSTAINABILITY AS PUBLIC POLICY

Despite its origins in ecology and the environmental movement, policymakers have embraced sustainability as a means of conceptualizing a broad range of principles and practices to guide decisions and actions in the present that do not disadvantage future generations or produce other negative externalities. In this sense, sustainability acts as a counterbalance to wellbeing, which often places the interests of those served in conflict with their competitors or successors (Kuhlman, 2010). As such, sustainability has taken on economic, social, technological, and political dimensions as well as its traditional physical and environmental attributes (Magis, 2008). In this context, social sustainability is defined as the capacity of a community to develop and engage its resources to adapt and overcome challenges amidst intense change (Magis, 2010). Resources include all sources of community capital – natural, technical, financial, social, cultural, political, and human (Flora & Flora, 2008).

RESILIENCE: BEYOND RESPONSE AND RECOVERY

Conventional approaches to emergency management emphasize four phases: prevention/mitigation, preparedness/readiness, response, and recovery. The progress of each phase is usually defined by resource capacity. Resilience frameworks defined by community resources often overlook issues of individual ownership and the influence of collaborative problem solving on resource sharing and replenishment.

Government agencies often overlook private sector resources in planning for disasters. Meanwhile, private sector organizations often act in their own individual interests and may have little regard for the decisions and actions of others. In all too many instances, inefficient or ineffective responses result from incorrect assumptions about the capacity or availability of government resources to respond and inadequate consideration of the value of private sector actors leveraging their capabilities for interests of the communities in which they operate. The ability of organizations or communities to address resource challenges usually depends upon leaders' abilities to manage resource capabilities as well as capacity. In this context, management refers to the ability to:

- Assess — *will it work?*
- Access — *can we get it?*
- Employ — *can we use it?* and
- Evaluate — *did it work?*

The attributes of community resilience – common purpose, shared values, creativity, equity, and trust, among others – transcend the utility of specific resources in effective crisis response.

This distinction helps explain why seemingly similarly situated communities often experience disaster differently. Communities with apparently low levels of preparedness can still respond to and recover from disasters efficiently by reducing conflict through collective action and collaborative problem solving. Studies of the response and recovery following the Canterbury Earthquake Sequence (Stallard, 2013; Carlton, 2017) and Great Tohoku Earthquake (Aldrich, 2015; Ye, 2019) suggest social capital and its expression strongly influence the disaster response and recovery (Aldrich, 2017).

SENSEMAKING IN ORGANIZATIONS

Weick (2005) defines sensemaking as the ability of individuals and organizations to rationalize organizational structure and behavior (decisions) with the ambiguity and complexity of their environments. This work emphasizes the influence of this retrospective capability to influence and shape the capacity of individuals and organizations to cope with disruptive events. Maitlis (2005) explores the role and influence of leaders in guiding sensemaking within organizations. Conklin (2006) addresses the influence of fragmentation and social complexity on wicked problems and leaders' attempts to "tame" them. His approach to dialogue mapping, like appreciative inquiry (Cooperrider, 2001) and other values-based approaches to organizational development (Sutcliffe, 2003), frames these challenges in terms of positive attributes as opposed to the dominant deficit-based approach that emphasizes damages, loss, disruption, and dysfunction. These perspectives align well with the public policy and administration literature on the roles of public officials in creating public value (Bennington, 2011) and partnering with communities in the co-production of public services (Bovaird, 2015).

An Integrative Framework

The proposed approach to public engagement draws upon the sources and traditions cited above. The framework articulated below draws upon the experience of the primary author as a local government official responsible for forging and fostering constructive engagement among public and private entities to promote crisis response and disaster recovery capabilities. This experience included managing emergency management operations for a metropolitan U.S. city and bringing together diverse public and private stakeholders to develop comprehensive plans for all hazards, all risks, and all phases of the emergency management process.

The five attributes of the framework reflect an emerging and untested sense of the virtues required to produce and sustain resilient systems. These virtues reflect the observed traits of successful public and public-private partnerships employed by the primary author's former municipal employer to empower communities and engage constituencies in changes to promote sustainability. These initiatives included programs to promote transportation alternatives, walkable neighborhoods, local food production, land, air, and water stewardship, alternative energy production, green buildings, and safety net services for vulnerable citizens. The attributes described below recognize common characteristics of these programs and serve as organizing elements for a focused stakeholder engagement experience similar to an appreciative inquiry summit.

The following five elements look beyond resource capacity to define the capabilities within a given community to put resources to use before, during, and after disruptive events:

- **Open** – How can individuals and organizations share data and insights (intelligence) about resources they control with one another without compromising data integrity, personal privacy, or competitive advantage?

- **Varied** – What critical infrastructure pathways (channels), key resources (connections), and information sources (content) support confidence, reliability, and redundancy under crisis conditions?

- **Simple** – What systems or methods are available and used to manage complexity, resolve ambiguity, and support coherence across organizational boundaries? How do these approaches create or contribute to shared experience or a sense of common purpose?

- **Local** – What resources – natural, technological, financial, political, social, cultural, human – exist within the community to sustain life and maintain critical functions during extreme events? What decisions or actions will individuals and organizations take to fill identified gaps?

- **Connected** – How are individuals and organizations integrated within regional, national, and international networks and cooperative agreements that can provide critical insights or key resources before, during or after a crisis?

These virtues represent attributes of resilient systems that achieve the following leadership aims:

- **Accuracy** – is the approach right, fit for purpose?
- **Accessibility** – is it available?
- **Adaptability** – will it accomplish what we want? and
- **Accountability** – can we stand by it?

SUMMARY AND CONCLUSIONS

Community response and recovery amidst crisis requires the combined efforts of public and private actors. Co-production of public services assumes the alignment of interests and actions among public and private actors enables collective responses that ensure effectiveness, if not improved efficiency. The shared interests of public and private actors that enable co-production lie at the heart of public value. The proposed approach seeks to extend the development and application of sensemaking beyond short-term tactical applications to enable a more coherent and strategic context informed by identified interests, unique capabilities, and common purpose to support organizational and community resilience. In doing so, it seeks to shape relationships and foster partnerships within communities to facilitate adaptative and collaborate action to overcome common challenges.

Desired Outcomes

The proposed public engagement framework seeks to increase social capital in all its forms – bridging, bonding, and linking – by forging alignment and a sense of shared purpose among public and private leaders that looks beyond the resources they control, especially fiscal, natural, physical, and human capital. By focusing explicitly on how leaders interact with one another, engage problems, exchange information, and make decisions, the proposed framework seeks to affect the experience of disruption and disaster by developing a coherent and constructive (positive) framework for collective action and collaboration. This approach stands in contrast to the deficit-based approach that most often starts with the questions, "What have I lost or what do I need?" and asks instead, "What can I offer or how can I contribute?"

Workshop Aim

This workshop is an opportunity for experienced emergency management practitioners, public policy analysts, decision-support researchers, and crisis communication experts, among others, to expand, develop, and evaluate the proposed framework for public participation and community engagement. Participants in this workshop will contribute to an enhanced understanding of this framework, its applicability to various policy and practice challenges, and how it can be developed and applied as a tool for community engagement, comprehensive emergency planning, and policy analysis. Participants will be invited to collaborate in further development, deployment, and evaluation of this framework as a community engagement strategy.

REFERENCES

Aldrich, D.P. (2015). It's who you know: Factors driving recovery from Japan's 11 March 2011 disaster in *Selected Works of Daniel P. Aldrich*. Available online at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.903.4229&rep=rep1&type=pdf, last accessed 10 June 2021.

Aldrich, D.P. (2017). The importance of social capital in building community resilience in *Selected Works of Daniel P. Aldrich*. Available online at https://www.researchgate.net/profile/Daniel-Aldrich/publication/316117238_The_Importance_of_Social_Capital_in_Building_Community_Resilience/links/5aa925d3458515178818a3f9/The-Importance-of-Social-Capital-in-Building-Community-Resilience.pdf, last accessed 10 June 2021.

Australian Public Services Commission (2007). Tackling wicked problems: A public policy perspective. Available online at https://legacy.apsc.gov.au/tackling-wicked-problems-public-policy-perspective, last accessed 10 June 2021.

Bennington, J. & Moore, M.H. (2011). *Public Value: Theory and Practice*. Palgrave MacMillan.

Bovaird, A.; Van Ryzin, G.G.; Loeffler, E. and Parrado, S. (2014). Activating citizens to participate in collective co-production of public services. Cambridge University Press. Available online at https://core.ac.uk/download/pdf/185481117.pdf, past accessed 10 June 2021.

Carlton, S. and Vallance, S. (2017). The commons of the tragedy: Temporary use and social capital in Christchurch's earthquake-damaged central city. *Social Forces* 96(2), pp. 831-850. Available online at https://academic.oup.com/sf/article-abstract/96/2/831/4107208, last accessed 10 June 2021.

Churchman, C.W. (1967). Guest Editorial: Wicked Problems. *Management Science* 14(4), pp. B141-142.

Conklin, J. (2006). Wicked problems and social complexity. CogNexus Institute. Available online at http://qi.speedwaystaging.co.uk/wp-content/uploads/2015/08/cn-wickedproblems.pdf, last accessed 10 June 2021.

Cooperrider, D.L. and Srivastva, S. (1987). Appreciative inquiry in organizational life. *Organizational Change and Development*, Vol. 1, pp. 129-169. Available online at https://www.researchgate.net/publication/265225217_Appreciative_Inquiry_in_Organizational_Life, last accessed 10 June 2021.

Flora, C.B. and Flora, J. (2008). *Rural Communities: Legacy and Change*, 3rd ed. San Francisco: Westview Press.

Kuhlman, T. and Farrington, J. (2010). What is sustainability? *Sustainability* 2, pp. 3436-3448. Available online at https://www.mdpi.com/2071-1050/2/11/3436/pdf, last accessed 10 June 2021.

Magis, K. & Shinn, C. (2008). Emergent Principles of Social Sustainability in Dillard, J., et al, ed., *Understanding the Social Dimension of Sustainability*, 1st ed. Routledge, pp. 15-43.

Magis, K. (2010). Community resilience: An indicator of social sustainability. *Society and Natural Resources* 23(5), pp. 401-416.

Maitlis, S. (2005). The social processes of organizational sensemaking. *Academy of Management Journal* 48(1), pp. 21-49.

Rittel, H.W.J. and Webber, M.W. (1973). Dilemmas in the general theory of planning. *Policy Sciences* 4, pp. 155-169.

Stallard, K. (2017). Neighbours and Social Capital in the Wake of the Christchurch Earthquakes. University of Otago, Dunedin, NZ. Available online at https://ourarchive.otago.ac.nz/handle/10523/3930, last accessed on 10 June 2021.

Sutcliffe, K.M. and Vogus, T.J. (2003). Organizing for resilience in Cameron, K.S., Dutton, J.E. and Quinn, R.E., eds., *Positive Organizational Scholarship: Foundations of a New Discipline*. San Francisco: Berrett-Koehler.

Weick, K.E.; Sutcliffe, K.M.; and Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science* 16(4), pp. 409-421. Available online at https://www.researchgate.net/publication/211395920_Organizing_and_the_Process_of_Sensemaking, last accessed 10 June 2021.

Ye, M. and Aldrich, D.P. (2019). Substitute of complement? How social capital, age and socioeconomic status interacted to impact mortality in Japan's 3/11 tsunami. *Population Health* 7. Available online at https://www.sciencedirect.com/science/article/pii/S2352827319300308, last accessed 10 June 2021.

Panel: What Good Are Models?

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Co-Chairs: Carrie REALE[a], Christen SUSHEREBA[b],
Participants: Shilo ANDERS,[a] David GABA,[c] Amanda BURDEN,[d] Laura MILITELLO,[b]
[a]*Vanderbilt University Medical Center, Nashville, TN*
[b]*Applied Decision Science, LLC, Unveil, LLC, Dayton, OH*
[c]*Stanford School of Medicine, VA Palo Alto Healthcare System, Stanford, CA*
[d]*Cooper Medical School of Rowan University, Camden, NJ*

A panel of experts will lead an exploration of the question, *what can we do with models of decision making?* Panelists will present their experiences applying naturalistic decision-making models in research and practice.

Topics explored:

- Decision making models related to healthcare
- The use of models to improve patient safety
- The use of models as a lens to interpret patient safety events
- The use of models to guide the design of recognition skills training

The Landscape of Decision Models

Shilo Anders, PhD is an Associate Professor in Anesthesiology and Biomedical Informatics at the Vanderbilt University Medical Center and Computer Science at Vanderbilt University, and a member of the Center for Research and Innovation in Systems Safety (CRISS). Her research applies cognitive systems engineering and health services research to medical contexts, specifically to improve the safety, performance, and satisfaction of humans working with technology and the built environment. Dr. Anders studies workflows and decision making of providers and patients for the design and development of systems.

This talk will summarize and synthesize naturalistic models of decision making, highlighting how models are both synergistic and separate. I will also highlight a hybrid model of decision making that builds on previous decision making models and incorporates common errors. I will discuss how the hybrid model was used to inform the design of a cognitive task analysis interview guide as well as the subsequent data analysis coding scheme.

Inside-out/Outside-in approach

David Gaba, MD is a Professor of Anesthesiology at Stanford School of Medicine and a Staff Physician at VA Palo Alto Healthcare System. Dr. Gaba has conducted simulation-based research continuously since 1989 to understand decision-making and action of anesthesiologists managing challenging acute situations. His clinical and teaching experience plus the various study results and collaboration with colleagues led him to create and publish cognitive models of anesthesiologists' work, and to implement novel approaches in simulation-based training linking dynamic decision-making with CRM-inspired teamwork and team management.

In this panel I will address why naturalistic research addressing anesthesiologist decision making and acting is important and what roles it might serve in improving patient safety. I will describe the "inside-out"/"outside-in" approach that has allowed cognition-savvy clinicians to link up effectively with cognitive scientists to plumb the depths of real-time work in complex dynamic life-critical medical settings. I will also discuss what I originally expected models might be good for and how the complexity of human behavior and the results of recent and ongoing (much larger) studies has modified these original expectations and research goals.

Decision-making pitfalls in practice

Dr. Amanda Burden is Editor-in-Chief of American Society of Anesthesiologists Simulation-Based Education Editorial Board and Professor of Anesthesiology at Cooper Medical School of Rowan University. Her research and publications address the use of simulation to explore a range of physician education, teamwork, healthcare quality, and patient safety issues, and received an Anesthesia Patient Safety Foundation (APSF) Safety Scientist Career Development Award as well as funding from the Agency for Healthcare Research and Quality.

While anesthesiology as a field is considered to have a greatly improved safety profile, errors in care and harm to patients persist. I will explore these issues through a discussion of approaches to improve understanding of decision-making strategies and pitfalls using a case-based approach. I will present real patient cases from the Anesthesiology Closed Claims Project. I will then use decision-making models to review challenges and errors in patient care and clinician judgment during anesthesia practice. The Anesthesia Closed Claims Project was initiated to analyze closed malpractice claims and improve patient safety. The goal of the Anesthesia Closed Claims Project is to identify and disseminate through the academic literature, educational programs, and assessment vehicles, major safety concerns, patterns of injury and strategies for prevention for anesthesiologists.

Implications of NDM models for training design

Laura Militello is co-founder and Chief Executive Officer at Applied Decision Science, LLC, a research and development company that studies decision making in complex environments. She is also co-founder of and Chief Scientist at Unveil, LLC, a company that delivers recognition skills training to combat medics, emergency responders and others. Ms. Militello contributed to the development of a set of applied cognitive task analysis methods for use by practitioners, and coauthored a text-book on cognitive task analysis. She has published widely on the use of decision-centered design.

Naturalistic decision making (NDM) models such as recognition-primed decision making (Klein, Calderwood, Clinton-Cirocco, 2010) and the data frame model of sensemaking (Klein, Moon, Hoffman, 2006) describe how people assess, interpret cues, and act in complex environments. These models have important implications for creating effective recognition skills training to support people who must quickly size up a situation and act (Klein & Wolf, 1995). Recent advances in augmented reality technologies enable affordable, flexible presentation of photorealistic cues, a core component of recognition skills training. As training technology designers integrate augmented reality, NDM models provide valuable guidance for training design that leverages the engaging nature of augmented reality and also promotes perceptual skill and mental model development. I will discuss nine guidelines for leveraging augmented reality in creating effective recognition skills training that are derived from NDM models.

This expert panel discussion will encourage participants to challenge their assumptions about the purpose and goals of decision-making models and explore potential new areas for practical application of the theoretical concepts expressed through these models for improving patient safety and training. The goal of the session is to share how decision-making models have been used in practical contexts, such as conducting research, training, and understanding patient safety events.

We hope the discussion helps spark new ideas for how theoretical models may be applied in practice and for future research in healthcare decision making and other domains. Whether it is through developing or refining models, applying models to academic research, or considering how these models might enhance training curricula, anyone interested in furthering the usefulness of decision-making models in the real world would benefit from this session.

References

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, *4*(3), 186-209.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent systems*, *21*(5), 88-92.

Klein, G., & Wolf, S. (1995, October). Decision-centered training. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 39, No. 20, pp. 1249-1252). Sage CA: Los Angeles, CA: SAGE Publications.

Naturalistic Decision Making & Resilience Engineering Symposium

Workshop Proposal: Tactical Decision Games

Presenter: John F. Schmitt, ShadowBox, LLC

1. Overview. We propose to demonstrate the use of tactical decision games to train tactical decision-making under conditions of uncertainty, complexity, time pressure, shifting goals and other pressures.

2. Plan of Action.

   a. Topic: Tactical decision making in the military and law enforcement.
   b. Method: Tactical decision game (TDG) workshops. Participants will participate in an interactive, online TDG to exercise decision making under naturalistic conditions. Students will be presented a scenario requiring a decision. Each scenario consists of an oral description and a diagram. Students will be required to come up with a decision in the form of the orders they would issue to subordinates and will then participate in a guided cognitive critique of the decisions. The session will conclude with a discussion of the methodology. We propose to run at least one military session and one law enforcement session. Time limit: 90 minutes each. Participant limit: 20.
   c. Facilitator: John F. Schmitt is a research associate at ShadowBox, LLC, and a former U.S. Marine infantry officer. He pioneered TDGs in the Marine Corps starting in the 1990s and is the author of *Mastering Tactics: A Tactical Decision Games Workbook*. C.J. Douglas is a retired New York State Trooper and a colonel in the U.S. Marine Corps.

3. Who could benefit. Anybody interested in training cognitive skills in domains characterized by uncertainty; time pressure; unclear, changing or changing goals; physical danger; and high levels of tacit knowledge.

4. Outcomes. (1) Promulgate a training methodology. (2) Stimulate and shape future research into training cognitive skills in the military and law enforcement.

5. Relevance. This training approach has applicability to a wide range of domains that include the characteristics described above and are of high importance to society.

Contact:
John F. Schmitt
john@schmittzu.com
217-649-8806

Naturalistic Decision Making & Resilience Engineering Symposium

Workshop Proposal: Health Care Cue-Detection Exercise

Presenters: Joseph Borders, ShadowBox, LLC; Regina Urban, School of Nursing and Health Innovation, University of Texas at Arlington; Yan Xiao, School of Nursing and Health Innovation, University of Texas at Arlington

1. Overview. We propose to demonstrate a cue-detect video decision-making exercise on emergency cardiac care. The cue-detect video decision-making exercise illustrates a strategy for training learners to recognize and interpret critical cues in a dynamic healthcare setting.

2. Plan of Action.

   a. Topic: Critical cue detection in cardiac care.
   b. Method: A Cue-Detect ShadowBox exercise. Participants will participate in an interactive, online video exercise to train critical cue-detection skills in a cardiac care scenario. Participants will watch a training scenario video together to detect critical cues in patient condition and medical team performance. After making their own observations, participants will then be exposed to the cues that a panel of expert nurses were able to detect in the same video. Through a guided discussion, participants will then reflect what they gained from being exposed to the expert insights. The session will conclude with a discussion of the methodology and the results of a study conducted at the University of Texas at Arlington's School of Nursing and Health Innovation. Time limit: 60 minutes. Participant limit: 30.
   c. Facilitator: Joseph Borders is a research associate at ShadowBox, LLC. Regina Urban is a professor and nurse scientist at UT-Arlington's School of Nursing and Health Innovation. Yan Xiao is a professor and researcher at UT-Arlington's School of Nursing and Health Innovation.

3. Who could benefit. Anybody interested in training cognitive skills in domains characterized by uncertainty; time pressure; unclear, changing, or competing goals; and high levels of tacit knowledge.

4. Outcomes. (1) Promulgate a training methodology. (2) Stimulate and shape future research into training cognitive skills in health care.

5. Relevance. This training approach has applicability to a wide range of domains that include the characteristics described above and are of high importance to society.

Contact:
Joseph Borders
jborders@shadowboxtraining.com
937-631-6774

Interactive and Integrative Sessions

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Session Co-Chairs: Emily Newsome[a] & Laura Militello[b]
[a]*ShadowBox LLC*
[b]*Applied Decision Science LLC*

1. Two sentence teaser that succinctly communicates why someone should attend and participate in this session

- Frontline child welfare practitioners make decisions that affect the safety and well-being of children under conditions in which facts are incomplete and information is frequently unreliable.
- This panel will bring together experienced practitioners, researchers, and thought leaders in child welfare to explore safety assessments and decisions, how practitioners and systems have responded to the COVID-19 pandemic, and ways to improve systems in which these decisions are made.

2. Summarize your plan for the session

- **Summary**: We have assembled a panel of diverse perspectives including public child welfare administrators, researchers, and consultants in the United States and United Kingdom. Each panellist will provide a 7- to 10-minute presentation about pressing issues from their perspective.
- **Presentation topics:**
    - How has the COVID-19 pandemic changed how child welfare workers assess the safety of children? How has the pandemic changed other critical practices and decisions in child welfare? What have been some of the main adaptations and innovations that were implemented in response to the challenges posed by COVID-19?
    - How can child welfare systems improve safety outcomes on a systemic level? What are strategies for building a learning culture and moving away from a punitive approach to errors?
    - What are strategies for improving decision making in this complex, highly dynamic work system?
- **Format:** After panellists share their 7- to 10-minute presentations regarding the above topics, the session co-chairs will curate questions from the audience to encourage discourse about safety from a systems perspective. Co-chairs may also ask their own pre-prepared questions in order to stimulate cross-talk among panellists and to further the discussion in areas of commonality.
- **Confirmed panellists:**
    - Eileen Munro: Professor Emeritus of Social Policy, London School of Economics. Professor Munro is a leading researcher in risk assessment and decision making in child welfare. Her interests include how organisational cultures help or hinder good quality reasoning and practice as well as understanding the complex causal relationships, including organizational and social factors that can influence individual decision making.
    - Jan Flory: Consultant, ShadowBox LLC. Jan has extensive experience examining child welfare practice and decision making through various roles in public and private child welfare agencies, including in New York City and Ohio. Ms. Flory is a thought leader in adapting ShadowBox, an innovative cognitive skills training approach, for child welfare.
    - Tamara Chapman-Wagner: Deputy Director, Cuyahoga County (Ohio, USA) Division of Children and Family Services: Ms. Chapman-Wagner manages multiple units of child welfare staff in a large, urban, child welfare agency in the U.S. She is uniquely familiar with the challenges faced by her staff due to the COVID-19 pandemic, and how they have adapted to overcome adversity.
- **Candidate panellists (invited, but not confirmed):**
    - Representative from the Annie E. Casey Foundation: foundation that sponsors research, develops best practices, and promotes innovative solutions to improve child welfare.
    - Representative from Williams County, Ohio: small, rural child welfare agency in the U.S. with a unique perspective about the challenges posed by COVID-19 in rural areas.

3. WHO WOULD BENEFIT FROM PARTICIPATING IN THIS SESSION?

- Anyone who studies, participates in, or has interest in child welfare systems would find this session beneficial. Additionally, the panel will discuss issues that may be relevant to a broader audience, including topics such as training and promoting organizational resiliency.

4. HOW THE OUTCOME OF THIS SESSION MIGHT BE UTILIZED (I.E., WILL THIS SESSION RESULT IN A PAPER/BOOK, SHAPE FUTURE RESEARCH, PROMULGATE A VIEWPOINT/PERSPECTIVE/METHOD, ETC.)

- The discussion from this panel will disseminate important viewpoints from practitioners and researchers on emerging topics relevant to both the child welfare and NDM / REA communities. The panel and audience participants will explore opportunities for future innovation, which will lead to future research and/or practice innovations.

5. THE WIDER SCOPE OR RELEVANCE OF THIS SESSION

- Child welfare organizations are tasked with maintaining the safety of children under complex and ever-changing circumstances. As such, they must be highly adaptable and resilient to emerging challenges. Decisions made in child welfare agencies have important repercussions for children, families, and societies. Furthermore, the organizational complexities at play in child welfare agencies are generalizable to other high-consequence domains. Innovations for improving this specific system will be highly applicable to other areas in which life-altering decisions are made in ambiguous, fast-paced circumstances.

Advanced Automation Panel

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Session CoChairs: Laura G. Militello[a] & Emilie M. Roth[b]
[a]*Applied Decision Science, LLC*
[b]*Roth Cognitive Engineering*

1. TWO SENTENCE TEASER THAT SUCCINCTLY COMMUNICATES WHY SOMEONE SHOULD ATTEND AND PARTICIPATE IN THIS SESSION

- This session will bring together people from agencies that are currently developing, applying, studying, and regulating the use of the advanced automation in critical contexts such military and civilian aviation, medical devices, and power industry.
- If you are interested in design, development, and assessment of advanced automation in complex settings, this panel will bring together diverse perspectives and domains for a lively discussion of current and emerging issues related to human-automation teams.

2. SUMMARIZE YOUR PLAN FOR THE SESSION

Summary: We plan to assemble panellists from a range of U.S. government agencies to share the issues and challenges they face in developing, applying, studying and regulating advanced automation, artificial intelligence, machine learning in complex settings. These technologies are proposed for use in military operations, civilian aviation, railways, patient and physician-facing medical devices, and many others. While some themes are common across contexts, challenges vary depending on context, intended use, and intended user. Panellists will discuss the challenges related to integrating advanced automation in their domains, risks they anticipate, and human capabilities that should be preserved and supported.

We have invited 6 panellists. Some have expressed interest, but are unable to commit until they have completed a lengthy approval process within their organizations. Confirmed panellists include:

- DR. JESSIE CHEN is a Senior Research Scientist (ST) for Soldier Performance in Socio-Technical Systems with U.S. Army Research Laboratory, located in Aberdeen Proving Ground, MD. Her research interests include human-autonomy teaming, agent transparency, human-robot interaction, and human supervisory control. Dr. Chen is a co-chair of the International Conference on Virtual, Augmented, and Mixed Reality (under the International Conference on Human-Computer Interaction), and she serves as an associate editor for IEEE Transactions on Human-Machine Systems and IEEE Robotics & Automation – Letters. She guest-edited a special issue on “Agent and System Transparency” for IEEE Transactions on Human-Machine Systems (2020) and a special issue on “Human-Autonomy Teaming” for Theoretical Issues in Ergonomics Science (2018).
- PAUL SCHUTTE is a principal researcher in the Applied Cognitive Science Department at Sandia National Labs. He is working on a methodology for human-machine teaming that promotes collaboration between human and machines. He is applying this methodology to mission planning and to intelligence gathering. Previously, he was an engineering research psychologist for the AeroFlight Dynamics and the Aviation Applied Technology Directorates (AFDD/AATD) in the US Army Aviation and Missile Research, Development, and Engineering Center (AMRDEC). There he developed a human-machine function allocation research program called Optimally Crewed Vehicles as well as designing interfaces for controlling multiple UAVs. Prior to the Army, he spent 35 years as a researcher in human factors and human-computer interaction at NASA. He has developed AI software for detecting system failures. He has led programs in Flight Deck Design and Integration under NASA’s High Speed Civil Transport Research (HSR), Error Proof Flight Deck Design, and Single Pilot Operations. He developed a clean-slate, flight deck design called the Synergistic Allocation of Flight Expertise flight deck (SAFEdeck). He has a BS in Math (Mary Washington College), an MS in Computer Science (William and Mary), and an MS in Experimental Psychology (Virginia Commonwealth University).
- JORDAN MULTER is a principal Technical Advisor at the U.S. Department of Transportation’s Volpe Center where he manages the rail human factors program. This program supports human factors

research for the Federal Railroad Administration and the Federal Transit Administration. Dr. Multer joined the Volpe Center in 1989 began studying the impact of information design in the use of aviation charts. For the past 28 years, he has directed his research efforts examining the use of technology and automation in support of safer rail and transit operations. In addition to performing human-in-the-loop research in the use of new technology for use in trains and rail traffic management systems he conducted extensive field research learning how humans and technology interact under real world conditions. He has coauthored over 40 publications including technical reports, journal articles, conference papers and a book chapter. Dr. Multer earned his masters and doctorate in experimental psychology from the University of Connecticut.

Panellists pending organization approval include:

- A representative from U.S. Army Futures Command
- A representative from the U.S. Federal Aviation Association
- A representative from the U.S. Food and Drug Administration

Interactive activities/exercises: Panellists will be asked to prepare responses to three questions:

1. What are the biggest human-machine teaming challenges you face in integrating advanced automation, artificial intelligence, and machine learning?
2. What are the most concerning risks related to advanced automation, artificial intelligence, and machine learning in the next 15 years?
3. When you think of the end user in your domain, what human capability(ies) do you think it is critical that we continue to leverage? What human capabilities do we want to design to support, or at the very least avoid disrupting or disabling?

After panellists respond to each question, participants will be invited to join the conversation with comments and questions.

3. WHO WOULD BENEFIT FROM PARTICIPATING IN THIS SESSION?

Anyone who studies, designs, develops, or is part of a human-machine team would benefit from this session.

4. HOW THE OUTCOME OF THIS SESSION MIGHT BE UTILIZED (I.E., WILL THIS SESSION RESULT IN A PAPER/BOOK, SHAPE FUTURE RESEARCH, PROMULGATE A VIEWPOINT/PERSPECTIVE/METHOD, ETC.)

The discussion from this panel will serve as the seeds for future research. Our hope is that the panellists who shape and fund research that informs the design, use, and regulation of advanced automation will share insights that inspire session participants. Another intended outcome of this session is that panelists will have an opportunity to think about and discuss their challenges through an NDM/REA lens.

5. THE WIDER SCOPE OR RELEVANCE OF THIS SESSION

As automation and other advanced technologies evolve, important decisions will be made about how and when to integrate them, the role of humans, and how best to regulate and introduce safety practices. These issues will touch nearly every aspect of our lives. As members of the NDM and Resilience Engineering communities it is important to engage in and help shape the discussion of how to build most effective joint human-automation systems maintain resilient performance – especially in the face of unanticipated variability.

PLEASE LIMIT YOUR SUBMISSION TO TWO PAGES OR LESS

Industry perspectives

Industry Partner

Extended Abstract Format for the Naturalistic Decision Making and Resilience Engineering Symposium

Thai Wood[a],

aResilience Roundup

1. Two sentence teaser that succinctly communicates why someone should attend and participate in this session

- The tech industry has rapidly grown into being a domain that touches many parts of our lives, meaning that it has the potential to have very high consequence incidents, while the increased speed of operations can create an ever increasing time pressure; at the same time the tech industry is historically one that is resistant to importing knowledge from some other disciplines and domains. In this session, I'll share the tactics and strategies that have worked for me in introducing resilience engineering concepts to teams at high tech companies.

2. Summarize your plan for the session

In the session, I'll draw upon my background as an engineer and incident responder in big tech corporations, a former EMT, and someone who studies resilience engineering and produces analyses for others. I'll discuss:

- Common challenges in tech, such as:
    - Resistance to importing knowledge from other disciplines
    - Relative newness of many companies and the industry as a whole
    - The possibility of high consequences at very fast speeds, coupled with little regulation or guidance
    - Encouragement to eschew traditional approaches
- Strategies that have been successful such as:
    - Introducing the ideas of a sociotechnical system
    - Exploring ways that automation can fail
    - Building off of increasingly accepted work such as blameless approaches to incident learning
    - Common ground in incident response
- I'll also explore subjects that tend to meet more resistance including:
    - That increased automation doesn't necessarily mean decreased incidents or accidents
    - Automation as a team player doesn't mean waiting for increased AI

3. Who would benefit from participating in this session

This session is relevant to anyone who works with or in tech companies or other mostly digital industries, also those that work with newer organizations that may be more resistant to change or adopting approaches that they view as more "academic".

4. What outcomes would you like to see as a result of this discussion?

This discussion should help others who are interested in introducing resilience engineering to their organizations as well as sparking discussions around ways of introducing topics and overcoming obstacles to adoption.

I hope this discussion will serve as a base for others to compare and contrast their experiences so that I and others can learn from other approaches as well.

5. The wider scope or relevance of this session

As software continues to expand to touch almost every aspect of our lives, it is critically important that the methods and learnings of safety sciences make their way into the digital domain. How, we as practitioners, introduce this knowledge and approach this subject can play a huge role in whether or not organizations take notice and explore sources of resilience internally or not. Due to this widespread nature, rarely if ever have we had the opportunity to increase the safety of such a large group of people or a chance to introduce these concepts to such a wide audience.

Citizen Audits for Resilience and Preparedness (CARP)

Ian RODERICK[a], May WHEELER[b] and Simon GILL[c]
[a] The Schumacher Institute
[b] Sustainable Food Trust
[c] The Safeguarding Community

1. Two sentence teaser that succinctly communicates why someone should participate and use your work

The idea of citizen audits for resilience and preparedness is presented for critical discussion. This could be a democratic evaluation and mechanism to engage communities in the resilience agenda and meet top down approaches with grassroot knowledge and opinion.

2. The Research

When facing shocks and stresses we aim for communities to not just bounce-back and survive, but bounce-forward and thrive. The principle in place has been to enable 'preparedness' to facilitate effective planning and response and 'resilience' to support anticipation, monitoring, response and learning. However, the coronavirus pandemic has exposed deep problems in how we approach resilience and preparedness at a community level.

Deliberative democratic processes have risen in popularity across the world in the form of citizen assemblies. Such assemblies have been successful and have been applied in areas such as abortion and gay marriage legislation (the classic Irish example) but also issues in transport, social policy, political appointments, and electoral reform. These assemblies work in parallel to formal governance mechanisms and act as guidance for decision making that is based on consensus among a large group of citizens drawn at random from the community.

Social audits are a form of citizen participation that looks at government performance and accountability. The objective of a social audit is to monitor and evaluate government performance - making public officials accountable for their actions and decisions. A social audit exercise is a mechanism of oversight: that is, the control that citizens can exert on their government officials to ensure that they act transparently, responsibly, and effectively.

This paper outlines a study of citizen assemblies and social audits and explores an application called a citizen audit for resilience and preparedness (CARP). This evaluates the capabilities and capacities of a community to respond to stresses and shocks, the process overseen and performed by a randomly selected group of citizens. Resilience and preparedness are characteristics that should emerge from a community rather than something imposed on it, a citizen audit could go some way towards building an active relationship between government bodies and its community.

Also discussed is a means of measuring success of the initiative, based on the regular reviewing of individual well-being, which can be run before and after a disruption.

3. WHO WOULD BENEFIT FROM KNOWING ABOUT YOUR WORK

At this stage the research is a Gedankenexperiment, (German: “thought experiment”) that explores possibilities and tries to think through what would be needed to setup such an audit process and what the consequences might be. It is an invitation to contribute ideas.

4. THE IMPACTS OR IMPLICATIONS OF THE RESULTS

Extending the benefits of citizens assemblies to preparedness and resilience has the potential to address the imbalances seen through the lens of Coronavirus. We must help communities to be better prepared for the next shock but also the creeping stressors of environmental change.

Papers

Uncovering resilience abilities in maintenance teams for buildings with Functional Resonance Analysis Method

Ivenio DE SOUZA[a], Assed HADDAD[a], Riccardo PATRIARCA[b]
[a]Programa de Engenharia Ambiental, Universidade Federal do Rio de Janeiro, Brazil
[b]Department of Mechanical and Aerospace Eng., Sapienza University of Rome, Italy

ABSTRACT
Building maintenance comprises a set of complex socio-technical activities, with many interacting agents. Traditional tools for safety management make the results of such evaluations distant from real situations. Resilience Engineering (RE) argues that improvements in safety performance concern the ability to recognize and adapt to handle unanticipated perturbations. This study aims at increasing the understanding of everyday building maintenance activities for air conditioning systems to uncover resilience abilities developed by the maintenance team to deal with everyday challenges. The Functional Resonance Analysis Method (FRAM) has been applied to model these activities. The data collection comprised interviews and observations. As a contribution, this study outlined the potential of the FRAM model as the basis of an in-depth and systematic analysis of daily performance, highlighting resilience abilities aligned with RE principles.

KEYWORDS
Resilience abilities; FRAM; building maintenance.

INTRODUCTION

The air conditioning systems constitute crucial types of equipment in building infrastructure, as responsible for maintaining good indoor air quality through adequate ventilation with filtration and providing thermal comfort for the building's occupants (Antoniadou & Papadopoulos, 2017). Therefore, adequate maintenance is essential to maintain air conditioning systems running and prevent any hazardous failure that can bring risk to the building's occupants.

In the building environment, the working conditions can all pose challenges to managing safety (Oswald et al., 2018). Also, the workers are exposed to hazards difficult to measure, since the ever-changing workplaces may potentially affect all workers on the site (Rosa et al., 2015). In addition, the absence of a preventive maintenance strategy increases the complexity of the work, once workers need to deal with scarce resources, poor tools, and insufficient teams to cover all the buildings on the university campus. There is a risk of occupational injuries due to the heavy load and physical demands of the activity. Often, mechanics deal with downgraded sites, confined spaces, and hard-to-reach places, especially when the air conditioning devices are installed in the roof or the underground. It is hard to predict or expect how the work will occur, given that work is done daily in different scenarios. Different work situations and emergency scenarios require a variety of responses that prescribed guidelines and procedures are sometimes unable to predict. In this sense, human performance can be understood as flexible and inherently variable (Wahl et al., 2020) that could be allied to cope with demands that complex socio-technical systems require daily.

The traditional safety concept referred to as the Safety-I presumes that things go wrong because of identifiable failures or malfunctions of technological components, procedures, beyond the humans, acting alone or collectively (Hollnagel et al., 2015). This approach is most useful in a system consisting of purely technical elements (Ham, 2020), once assumes it is always possible to identify a linear dichotomic cause/effect relation (Patriarca, Falegnami, et al., 2018) and removing or weakening the causes of adverse outcomes can improve safety (Hirose & Sawaragi, 2020). However, traditional tools in line with Safety-I vision are insufficient to provide a complete and comprehensive representation of the work-as-done (WAD). This stems from the fact that reality is complex, variable, and even unpredictable and working conditions are rarely ideal (Pardo-Ferreira et al., 2020). Such characteristics of Safety-I indicate that the respective approaches should not be used to improve safety in work environments where workers at the sharp end have established safety practices that pervade work activities themselves (Saldanha et al., 2020).

Conversely, the Safety-II vision argues that workers play a significant role in safety management, once human flexibility, and their ability to adjust work to deal with varying conditions instead of strictly following operational rules could contribute to systems working correctly (Lee et al., 2019). Therefore, the safety-II vision encourages a greater emphasis on the aspects that contribute to normal performance (Harvey et al., 2019).

Thus, for reliable analysis and a better understanding of these systems, it becomes necessary to apply a perspective in line with Safety-II and Resilience Engineering (RE), which concerns a new approach for safety management that focuses on how systems anticipate undesirable conditions or managing changes, and thereby to continue the operation, even after a disruptive event, or the presence of continuous stress (Hollnagel et al., 2006).
The resilience abilities have contributed to creating a wide consensus on resilience structure, in which the resilience comprises what the system does rather than something the system has (Patriarca, Bergström, et al., 2018). Hollnagel (2017) defines these abilities as i) monitoring, which concerns developing abilities to cope with near-term events and to monitor what happens in the operating environment; ii) responding, which consists of actions to be adopted to appropriately respond to changes in the system environment; iii) learning, which means an organization modifies or acquire new knowledge, competencies, and skills on everyday work and activities; and iv) anticipating, which lies in attempting to prepare for further events such as disturbances or improvements in the system functioning.
The present study aims to identify emergent resilience abilities in the maintenance activities in line with the four core abilities of resilient systems proposed by Hollnagel (2017). The empirical field of study is a university campus in Rio de Janeiro, Brazil. The analysis was described using the Functional Resonance Analysis Method (FRAM) (Hollnagel, 2012) and outcomes were analysed from the perspective of RE, that is, how the mechanics manage variabilities and disturbances to achieve successful maintenance in their everyday work.

METHOD

The method applied in this research is the Functional Resonance Analysis Method (FRAM) (Hollnagel, 2012), which has been developed under the concepts and principles of the RE. FRAM comprises methodological support for modelling the varieties of the work domain under investigation (Patriarca et al., 2019). Figure 1 presents the framework for analysing empirical resilience abilities, which has three stages.

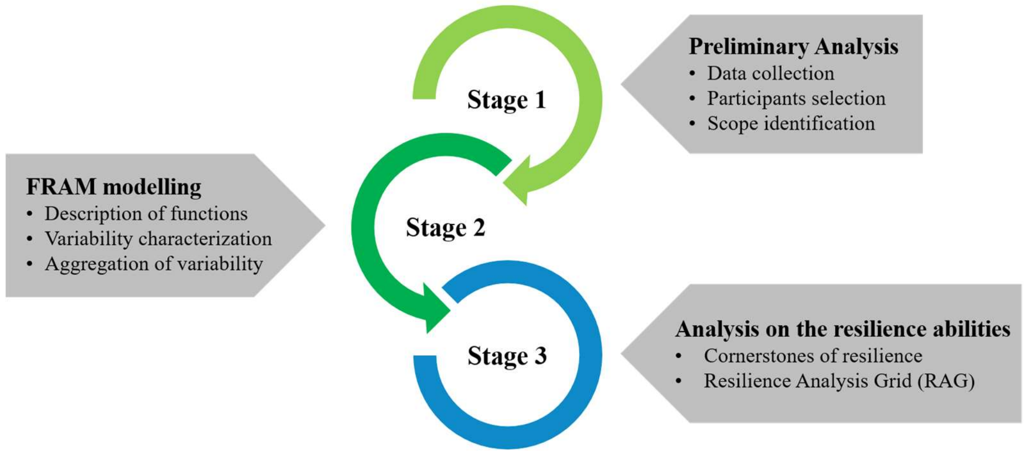


Figure 1. Proposed framework for analysing empirical resilience abilities

Stage 1 is concerned with establishing a preliminary analysis on the common ground where the research was performed. This stage included the scope identification as well as selecting the participants for data collection. The maintenance of air conditioning devices was selected for the study. The main reason for selecting this subject was that it had been part of a recent study (Souza et al., 2021), which made it easier to access for data collection.
Stage 2 is dedicated to a deeper understanding of WAD by using the FRAM. The following is a brief description of the four steps for developing a FRAM model (Hollnagel, 2012). The first step concerns the description of essential functions to perform an activity. These functions can be human, technological, or organizational, depends on its natures in the system. Each function is represented up to six aspects, consisting of one output and five inputs:

- Input (I): what trigger the function or what is processed or transformed by the function
- Output (O): what is the result of the function, it can be either a state change or a specific product
- Precondition (P): mandatory conditions that must exist before the function can be performed
- Resource (R): what the function needs or must consume when it is carried out to produce the result (the output)
- Control (C): what controls and monitors the function to match the desired output
- Time (T): temporal requirements or constraints of the function, regarding both duration and time of execution.

Once the function description is done, the second step is the identification of the output variability of each function of the model, characterising each function with its potential and actual performance variability. After this, a third step consists of examining instantiations of the model to understand how the variability of each function can be resonant in the system. The fourth and last step is the monitoring and managing of the performance variability of each proposed instantiation of the model.
Since the variability analysis in stage 2 is completed, a third and last stage is needed, which a set of questions

based on the idea of the Resilience Analysis Grid (RAG) (Hollnagel, 2011) had been used in attempting to categorise empirical abilities captured in each FRAM's function into the four cornerstones of resilience. The definition of general questions applicable to each domain is hardly possible, once resilience is strongly related to the system's purpose for which is being assessed (Patriarca, Di Gravio, et al., 2018). Therefore, it requires the analyst to adjust its structure to the domain-specific context under study (Chuang et al., 2020). Thus, in the current study open questions based on the RAG has been tailored to fit the context of the maintenance in air-conditioning devices.

Data collection and participants

Observations and interviews supported the data collection in the fieldwork. The procedures included direct observation from the work planning meeting of the maintenance team up phase of intervention in air conditioning devices. The open-ended approach had been chosen for the interviews with the participants, in which broad and open questions should be asked, and the replies to them should inform the researcher of the perceptions of the individuals (Sekaran & Bougie, 2016). The objective of the data collection was to gather information to make it easier to understand the tasks that would be subsequently modelled and analysed with FRAM.

Four maintenance workers from the university campus participated in the data collection. Three mechanics trained in split-type air conditioning, and the maintenance supervisor (civil engineer). In everyday activities, the mechanics are responsible for installs, maintains, and repairs of air conditioning devices for all buildings of the organization. The maintenance supervisor plays the role of receiving and analysing requests, issuing work orders (WOs), and offers technical support to the mechanics. The results of the analysis have been validated through semi-structured interviews with workers involved in the study.

RESULTS AND DISCUSSIONS

The purpose of this section is to explore the potential of applying the proposed framework to categorise empirical resilience abilities into the four resilience abilities. A case study is presented in the building maintenance domain, focusing on understanding the WAD in the maintenance of air conditioning devices.

Stage 1 – preliminary analysis

The case study was undertaken on a Brazilian university campus located in the city of Rio de Janeiro. The department for building maintenance is a facilities management unit responsible to maintain acceptable use conditions in the buildings of the campus. The building maintenance covers six areas: electrical, plumbing, air conditioning, civil works, metal works, and carpentry/furniture. The current study is focused on the maintenance of air conditioning devices. Other areas may be disclosed in upcoming studies.

The overall process of building maintenance comprises three major stages: maintenance request, request analysis, and maintenance execution. The current study focuses on the stage of maintenance execution for air conditioning devices. The study on the other stages can be reached in an earlier study (Souza et al., 2021).

Stage 2 – FRAM modelling

The FRAM model for the maintenance of air conditioning systems consists of sixteen functions, as illustrated in Figure 1. Workers who participated in the FRAM development suggested that only eight functions had significant variability: i) check requirements request; ii) choose the best route; iii) triage of WOs; iv) perform customer visit; v) access the site; vi) access air conditioning device; vii) perform repair; viii) register spare parts. Such functions are depicted in green and with a sine wave. Although the variability related to the outputs may be described by using multiple phenotypes, i.e., in terms of time/duration, force/distance/direction, wrong object, and sequence, in this study the two main phenotypes proposed by Hollnagel (2012), i.e., timing (on time, too late, too early, not at all) and precision (precise, acceptable, imprecise) are adopted as they are enough to describe most outcomes.

FMV software (Hill, 2019) allows the graphical display of information and provides useful features to check the completeness of the functions. The construction of a model based on analysis of the everyday work as well as the variability analysis enabled extraction resilience abilities that would contribute to overcoming disturbances throughout the activity.

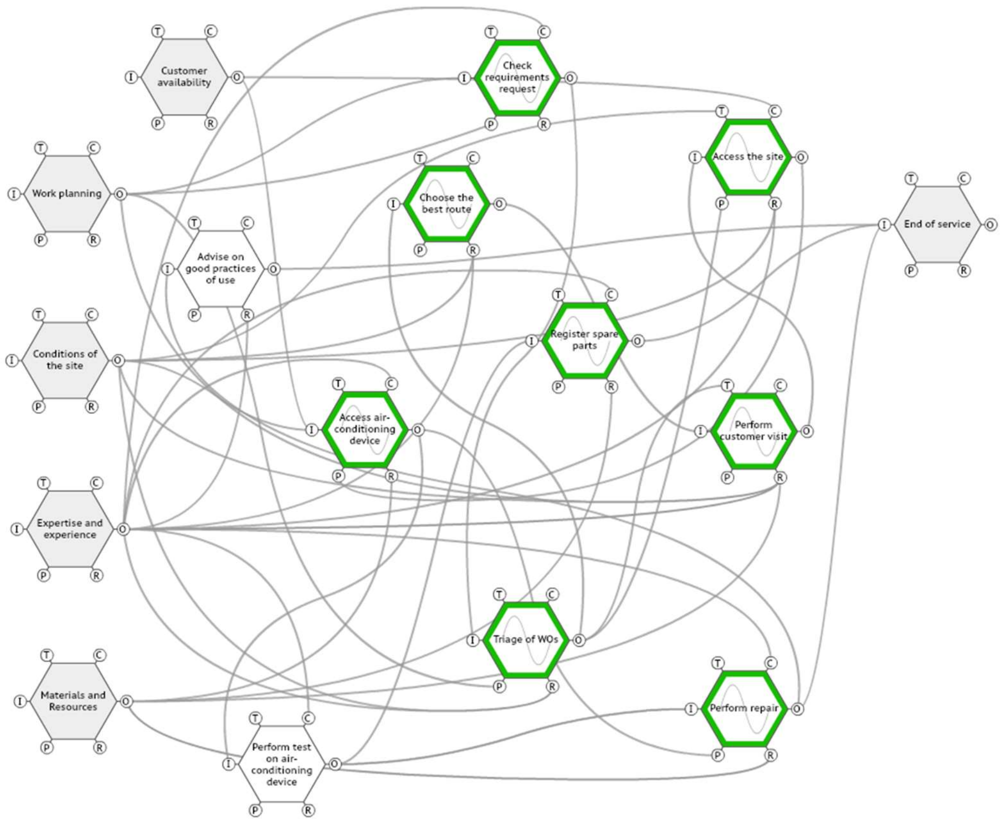


Figure 2. FRAM model for the maintenance of air conditioning devices. Functions depicted in green are functions presenting variability. The other functions depicted in white do not present variability

The function ‘work planning’ represents the planning for daily work. This function requires several decisions made in collaboration with the maintenance team concerning strategies to be adopted during the work shift. The issuance of WOs triggers the function ‘check requirements request’. This study does not intend to explore the issuance of WOs, however can be reached in previous work by Souza et al. (2021). The function ‘check requirements request’ consists of analysing WOs to find the requirements for the maintenance. In this function, mechanics should verify the service requested, the service location, and whether there are any specific demands. As indicated in Table 1, this function relies on the possibility of the check might not occur or might occur with reduced precision. In these cases, it affects the function ‘triage of WOs’, as priorities might be wrongly set.

The function ‘triage of WOs’ represents a cognitive process that happens every morning after meeting with the team. It uses variables mainly the work scope resulting from the function ‘check request requirement’, and resources such as knowledge of the local and time of route as decision-making to prioritize the WOs. As described in Table 1, this function is highly susceptible to variability in both timing and precision. Output quality depends on how thorough the failure description by the requestor is performed. Poor information entails the increase in the possibility of inadequate prioritization of WOs, which may affect the downstream function ‘choose the best route’.

The weather conditions act as a decision element to the function ‘access the site’, e.g., if there is a “heavy rain”, the workers do not perform external services. Conversely, on hot days the workers are exposed to rigorous solar radiation, causing high physical workload and fatigue. To mitigate these effects, when possible, they do not perform activities that expose them to solar radiation in time between 11 am and 3 pm. This setting may lead to variability regarding time faced weather conditions.

Also, functions ‘work planning’ and ‘materials and resources, support technically the performing of visits with WOs and working tools. The output variability regarding precision occurs when the WOs were not issued on time; therefore, the workers need to perform a visit without the WOs at hand.

The function ‘access air conditioning device’ acts as a precondition to the function ‘perform repair’. It also consumes some resources, such as equipment for access (e.g., ladder or scaffolds) because some outside units are installed on the roof or at height. Moreover, the absence of conservation in air conditioning structures is a contributing factor for accidents involving the maintenance team. There are a meaningful number of territories with air conditioning devices installed in places of difficult access or in confined spaces (i.e., devices installed on the roof or the underground) which significantly affect the workers’ performance.

The function ‘customer availability’ is also a temporal constraint to trigger this function because when the customer is unavailable workers cannot access the device to perform the repair. In this case, mechanics use their knowledge to identify near customers requiring maintenance to visit. The function ‘perform repair’ relies on the worker's expertise and experience, mainly regarding the knowledge they have on the specific device. Moreover, this function consumes several resources, such as working tools, an oxyacetylene torch, a vacuum pump, among others. Indeed, the number of resources depends on the work scope. In some cases, workers need to return to the workshop to get additional tools to perform the repair once WOs do not provide the work scope properly. As presented in Table 1, the potential variability regarding both timing and precision in the function 'perform repair' depends on spare parts availability and working tools.

The background function 'expertise and experience' controls the triggering of the 'register spare parts' function once the mechanics' expertise is crucial to perform the task. This function consumes resources as a standard form used by the mechanics to record the spare parts. As shown in Table 1, whereas output variability regarding time is "on time", variability regarding precision is "imprecise". The imprecise output results from the unavailability of the standard form to record spare parts. Mechanics need to appeal to the memory or a handwritten paper to record parts. However, these actions can lead to misunderstandings.

Table 1 summarizes the functions presenting variability as well as the characterization of output variability in terms of timing of precision. The interactions among functions had been graphically indicated in Figure 1, which shows the instantiation of the FRAM model for the analysed scenario.

Table 1. Variability analysis on the instantiation of the FRAM model

Function	Variability	
	Regarding time	Regarding precision
Check request requirements	Not at all If available information to work is poor, the output may not be performed at all.	Imprecise In the event of an incorrect or incomplete failure description by the customer, there may be an error in the analysis.
Triage of WOs	Too late This function is highly dependent on workers' expertise.	Imprecise Output quality depends on how thorough the requestor information is. Therefore, some WOs may be wrongly prioritized because of poor information from the requestor
Choose the best route	On-time This function comprises the decision-making of the team. It is a function relatively quick.	Acceptable Output precision depends on the workers' knowledge in the territory.
Perform customer visit	Not at all This function depends mainly on the weather conditions.	Imprecise Non-issuance of WOs can lead to execution error due to lack of information regarding work
Access the site	Not at all If the customer is not at the site or unavailable, the output may not be produced at all.	Acceptable Output precision depends on the workers' knowledge in the territory
Access air conditioning device	Not at all If the conditions for performing the maintenance are adverse, the output may not be performed at all.	Imprecise Quality of access to air conditioning devices depends on the conditions of the site (e.g., device installed in high outside position, confined space, or in the roof). These conditions may cause risky situations for workers or imply non-perform of the maintenance.
Perform repair	Not at all It depends on spare parts availability. If there is no spare part to perform repair, the output may not be performed at all.	Imprecise All the outputs depend on how thorough the tasks are performed.
Register spare parts	On-time Mental effort consumes relatively little time related to the activity.	Imprecise Unavailability of the form to record spare parts can lead to misunderstandings or errors.

Stage 3 – analysis on the emergence of resilience abilities

Table 2 presents the relationships between the functions, the resilience abilities that emerged from the field examination, and the four cornerstones of resilience. The four cornerstones were brought into this analysis in attempting to categorise the empirical abilities captured in each analysed function. From the variability analysis, a set of questions based on the idea of the RAG had been developed to determine how each empirical ability matches the four cornerstones of RE.

As shown in Table 2, the same empirical ability may be noticed in multiple functions. For instance, the knowledge of the territory and specifically the site where the work is done can be encountered simultaneously in more than three functions. Similarly, these empirical abilities seem to be associated with multiple core abilities depending on how they are employed.

The analysis disclosed that mechanics take advantage of their expertise and experience to interpret data from WOs. Also, when complete data are not available, they go to the site and check requirements *in-loco*.

The prioritization of WOs is supported by the workers' tacit knowledge. This decision-making relies mainly on the workers' knowledge about the territory they operate. The knowledge of the territory is a crucial element in delineating an optimized route to avoid unnecessary pathways.
The function 'perform customer visit' takes advantage of the expertise and experience of workers, such as workers' knowledge on the territory, and affinity with customers, this contributes to imitate an appointment. In this scenario, we could note that some WOs are issued verbally on the site by the mechanics. The supervisor reported that flexibility is vital to respond to unscheduled situations quickly, i.e., this variability enables the attending faster in specific demands (e.g., lack of energy and water leakage) or to optimize the route of workers. In this sense, at the operational level, resilience may be a function of how organizations deal with apparently contradictory requirements, i.e., good procedures and good plans are desired, while appropriate flexibility is encouraged to meet the real demands of the daily operation (McDonald, 2006).
The function 'access the site' is the reason for many complaints by the workers. The function 'customer availability' controls the decision-making about aborting or perform the repair. We noted some cases that on arrival at the site, the team did not find anyone to receive them. Faced with this situation, usually, they either visit another customer or return to the workshop. Also, the function 'access the site' uses resources like workers' expertise, i.e., the knowledge on specific conditions of the site to support the decision-making during the activity.
The situational context, e.g., the weather and site conditions are constantly changing and can be the source of small and big surprises (Siegel & Schraagen, 2017). Therefore, the ability to learn from past experiences and anticipate actions perform a crucial role in safe maintenance operations.
As aforementioned, the instant of accessing the air conditioning poses occupational risks to the maintenance team as well as exposes them to hazards. To overcome these constraints, the maintenance team makes sense on the current conditions to construct safety during their maintenance interventions. Thus, sensemaking (Weick et al., 2005) contributes to resilience since it focuses on sharp-end adaptations (Kilskar et al., 2018) for safer decisions, but also ways to have efficiency in the maintenance.
In providing a bill of materials, the procedures and guidelines provide that a standard form should be used for it. However, in the WAD this form is unavailable most of the time. When faced with it, the mechanics proposed to create a group in the messaging app to share relevant information about the WO. To perform the repair in the devices sometimes mechanics faces the unavailability of spare parts. However, workers adopt some strategies to cope with these constraints, as dismantling an old device to remove a part in-condition of use.
This study proposed an integrated framework that brings the four cornerstones of the RE into building maintenance practices. As previously stated, some questions inspired in the RAG were adopted, such as, does the maintenance team try to learn from failures (things that go wrong) as well as from successes (things that go right)? Also, how they cope with incomplete information? To check requirements on the WOs, sometimes the maintenance team deals with poor description for the failure in air conditioning devices. However, the mechanics have developed resilient abilities to deal with the specific demands of their activities, making use of the variability positively. For example, knowledge acquired from previous experiences contributes to interpreting data in WOs even though the available information has a high degree of subjectivity. Moreover, the workers' sensemaking on current conditions arises in the instant of accessing the air conditioning device; thus, they choose the better procedure to gain access to the device.

Table 2. Resilience abilities uncovered with the FRAM modelling

Function(s)	Description of resilience abilities	Cornerstones of resilience			
		Anticipating	Learning	Monitoring	Responding
Check requirements request	To check requirements in WOs, the mechanics engage their abilities to interpret data based on the past experiences		✓		✓
	Mechanics check *in-loco* the requirements for the maintenance when complete data has not been previously made available				✓
Work planning	Strategies to be adopted during the work shift are chosen in a daily work planning	✓			
Choose the best route / Triage of WOs / Access the site	The knowledge on the territory supports the decision-making in delineating an optimized route to visit the sites	✓	✓		✓
Perform customer visit	The relationship and affinity with the customers contribute to speeding up the service and overcoming limitations on missing information		✓		✓
	Workers monitoring the weather conditions to decide on the maintenance schedule			✓	
	Emergencies require WOs to be issued verbally on the site by the mechanics				✓

Access air conditioning device	Knowledge of the site and specific conditions supports the procedures to be adopted in the maintenance task		✓		
Perform repair	To overcome unavailable spare parts, mechanics maintain in the warehouse parts in-condition of use	✓	✓		
Register spare parts	Mechanics proposed to create a group in message app to facilitate sharing relevant information about the WOs when the fieldwork is in progress.			✓	

CONCLUSION

This article suggested a framework to identify empirical resilience abilities and check the adherence of them to the four cornerstones of resilience. This study used the FRAM to model routines in building maintenance for air conditioning of a Brazilian university campus.

The major finding in this study stem that the mechanics lack accurate information on current conditions to perform their activities. Nevertheless, knowledge of the territory seems to be a prerequisite for successful operations. Thus, the most visible manifestations of resilience in the maintenance activities are resulting of adaptability and sensemaking of the maintenance team.

The case study demonstrated that the FRAM model might offer empirical evidence for extracting resilience abilities from variability analysis. Moreover, questions based on the idea of RAG proved to be an effective means for categorising empirical resilience abilities into the four cornerstones of resilience. However, future studies are required to explore opportunities, such as i) examination of this framework in other domains, ii) a quantitative analysis of resilience from the FRAM modelling, and iii) developing of strategies and guidelines to enhance empirical resilience abilities.

ACKNOWLEDGEMENTS

The authors want to thank the Brazilian university, especially all the professionals involved in the maintenance of building air conditioning systems by willingness and collaboration in the data collection.

REFERENCES

Antoniadou, P., & Papadopoulos, A. M. (2017). Occupants' thermal comfort: State of the art and the prospects of personalized assessment in office buildings. *Energy and Buildings*, *153*, 136–149. https://doi.org/10.1016/j.enbuild.2017.08.001

Chuang, S., Ou, J. C., & Ma, H. P. (2020). Measurement of resilience potentials in emergency departments: Applications of a tailored resilience assessment grid. *Safety Science*, *121*(291), 385–393. https://doi.org/10.1016/j.ssci.2019.09.012

Ham, D. (2020). Safety-II and Resilience Engineering in a Nutshell : An Introductory Guide to Their Concepts and Methods. *Safety and Health at Work*, *xxxx*. https://doi.org/10.1016/j.shaw.2020.11.004

Harvey, E. J., Waterson, P., & Dainty, A. R. J. (2019). Applying HRO and resilience engineering to construction: Barriers and opportunities. *Safety Science*, *117*, 523–533. https://doi.org/10.1016/j.ssci.2016.08.019

Hill, R. (2019). FMV PRO 2.0.2. *13th FRAMily Meeting*. http://www.puroresu.com/wrestlers/londos/death.html

Hirose, T., & Sawaragi, T. (2020). Extended FRAM model based on cellular automaton to clarify complexity of socio-technical systems and improve their safety. *Safety Science*, *123*(June 2019), 104556. https://doi.org/10.1016/j.ssci.2019.104556

Hollnagel, E. (2011). Introduction to the Resilience Analysis Grid (RAG). In J. Pariès, J. Wreathall, & E. Hollnagel (Eds.), *Resilience Engineering in practice: A guidebook*. Ashgate.

Hollnagel, E. (2012). *FRAM: The Functional Resonance Analysis Method - Modelling the Complex. Socio-technical Systems*. Ashgate.

Hollnagel, E. (2017). *Safety-II in Practice: Developing the Resilience Potentials*. Routledge.

Hollnagel, E., Wears, R. L., & Braithwaite, J. (2015). *From Safety-I to Safety-II: a white Paper*. The Resilient Health Care Net: Published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia.

Hollnagel, E., Woods, D. D., & Leveson, N. (2006). *Resilience Engineering: Concepts and precepts*. Ashgate. https://doi.org/10.1201/9781315605685-1

Kilskar, S. S., Danielsen, B. E., & Johnsen, S. O. (2018). Sensemaking and resilience in safety-critical situations: A literature review. *Safety and Reliability - Safe Societies in a Changing World - Proceedings of the 28th*

International European Safety and Reliability Conference, ESREL 2018, 483–492. https://doi.org/10.1201/9781351174664-60

Lee, J., Yoon, W. C., & Chung, H. (2019). Formal or informal human collaboration approach to maritime safety using FRAM. *Cognition, Technology and Work*. https://doi.org/10.1007/s10111-019-00606-y

McDonald, N. (2006). Organisational Resilience and Industrial Risk. In E. Hollnagel, D. . Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 155–180). Ashgate Publishing, Ltd.

Oswald, D., Sherratt, F., & Smith, S. (2018). Problems with safety observation reporting: A construction industry case study. *Safety Science*, *107*(March), 35–45. https://doi.org/10.1016/j.ssci.2018.04.004

Pardo-Ferreira, M. del C., Rubio-Romero, J. C., Gibb, A., & Calero-Castro, S. (2020). Using functional resonance analysis method to understand construction activities for concrete structures. *Safety Science*, *128*(October 2019). https://doi.org/10.1016/j.ssci.2020.104771

Patriarca, R., Adriaensen, A., Putnam, J., Peters, M., Constantino, F., & Di Gravio, G. (2019). Receipt and dispatch of an aircraft: a functionalk risk analysis. *8th REA Symposium Embracing Resilience: Scaling up and Speeding Up*. https://doi.org/10.15626/rea8.17

Patriarca, R., Bergström, J., Di Gravio, G., & Costantino, F. (2018). Resilience engineering: Current status of the research and future challenges. *Safety Science*, *102*(102), 79–100. https://doi.org/10.1016/j.ssci.2017.10.005

Patriarca, R., Di Gravio, G., Costantino, F., Falegnami, A., & Bilotta, F. (2018). An Analytic Framework to Assess Organizational Resilience. *Safety and Health at Work*, *9*(3). https://doi.org/10.1016/j.shaw.2017.10.005

Patriarca, R., Falegnami, A., Costantino, F., & Bilotta, F. (2018). Resilience engineering for socio-technical risk analysis: Application in neuro-surgery. *Reliability Engineering and System Safety*, *180*(November 2017), 321–335. https://doi.org/10.1016/j.ress.2018.08.001

Rosa, L. V., Haddad, A. N., & de Carvalho, P. V. R. (2015). Assessing risk in sustainable construction using the Functional Resonance Analysis Method (FRAM). *Cognition, Technology and Work*, *17*(4), 559–573. https://doi.org/10.1007/s10111-015-0337-z

Saldanha, M. C. W., de Carvalho, R. J. M., Arcuri, R., Amorim, A. G., Vidal, M. C. R., & Carvalho, P. V. R. de. (2020). Understanding and improving safety in artisanal fishing: A safety-II approach in raft fishing. *Safety Science*, *122*(October 2019), 104522. https://doi.org/10.1016/j.ssci.2019.104522

Sekaran, U., & Bougie, R. (2016). *Research method for business: A Skill-Building Approach* (7th ed.). John Wiley & Sons. www.wileypluslearningspace.com

Siegel, A. W., & Schraagen, J. M. (2017). Team reflection makes resilience-related knowledge explicit through collaborative sensemaking: observation study at a rail post. *Cognition, Technology and Work*, *19*(1), 127–142. https://doi.org/10.1007/s10111-016-0400-4

Souza, I. T. De, Rosa, A. C., Vidal, M. C. R., Najjar, M. K., Hammad, A. W. A., & Haddad, A. N. (2021). Information Technologies in Complex Socio-Technical Systems Based on Functional Variability : A Case Study on HVAC Maintenance Work Orders. *Applied Sciences*, *11*.

Wahl, A., Kongsvik, T., & Antonsen, S. (2020). Balancing Safety I and Safety II : Learning to manage performance variability at sea using simulator-based training. *Reliability Engineering and System Safety*, *195*(September 2019), 106698. https://doi.org/10.1016/j.ress.2019.106698

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, *16*(4), 409–421. https://doi.org/10.1287/orsc.1050.0133

Development of a Debriefing Tool for Performance Evaluation in Maritime Training Simulations

Josué E. M. FRANÇA[a], Karolina STARK[a], Gesa PRAETORIUS[a,b] and Jan SNÖBERG[a]
[a]*Linnaeus University, Sweden*
[b]*University of South-Eastern Norway, Norway*

ABSTRACT
This article presents a proposal for a debriefing tool developed to support the development of bridge resource management skills of nautical science students. This debriefing tool consists of a set of questions that aim to trigger reflection about the bridge team's performance during simulator exercises. The tool has been tested by students in conjunctions with the ship handling exercises. After this test, feedback from the students has been obtained through a focus group. The results show constrains that may jeopardize the utilization of a debriefing tool, but also encourages discussions regarding undesirable and desirable outcomes, gathering a channel for feedbacks. The study also indorsed this tool as a way of enhancing performance through the understanding, development and training of individual competences needed for a safe onboard work.

KEYWORDS
Debriefing; Maritime Safety; Maritime Education and Training.

1 INTRODUCTION

Shipping can be analysed as a combination of technology (vessel structures, engine, equipment, instruments etc) and social system (the crew, their culture, norms, habits, custom, practices etc) (Grech et al., 2008). In this context, the workers can be seen as the element which connects technology, equipment and culture in a sociotechnical network of interactions that compose the maritime workplaces. Due to these and other characteristics, this industry is considered one of the high-risk domain with a high degree of complexity in operations (Le Coze, 2021). In this context, Maritime Education and Training (MET) is regulated through the International Convention on Standards of the International Maritime Organization (IMO). The standard formulates the minimum competence and certification requirements for seafarers, and the Standards of Training, Certification and Watchkeeping for Seafarers (STCW) is responsible for the requirements in training and certification of seafarers with a specific focus on the master and officers onboard. The aim of the convention is to establish the preconditions for comparable training standards globally (IMO, 2010).

These training standards includes technical and non-technical skills (NTS), providing the development of both by training the operation and maintenance onboard – technical skills, as well as communication and teamwork – non-technical skills. In fact, the updated revision of the STCW, in 2010, expanded the demands on NTS for officers onboard, by requiring proficiency in knowledge concerning leadership, management, and teamwork (IMO, 2010). This training, in addition to meeting the international requirements of these conventions, prepares seafarers for the daily demands of their work, involving their technical knowledge in what they do, associated with the necessary NTS. The integrated development of technical and non-technical skills is a necessity for this daily routine work, as well as for emergency occurrences that may happen. Seeking to learn and develop competences for these work situations during the training simulations, this research proposes a debriefing tool to help students to understand the skills needed for a safe and productive performance.

2 THEORICAL FRAME OF REFERENCE

This section provides the theoretical background for this research, introducing the concepts an evolution of the resource management training approaches, considering the NTS related to that, contextualized in the debriefing tool methodologies and applications.

2.1 Non-technical skills

Non-technical skills (NTS) can be defined as the cognitive, social and individual resources competences that complement technical skills, enabling a safe and productive worker performance (Flin et al., 2016), being eminently individual, since it is observed in the interaction of the worker with all parts that form the work environment. Having a broader perspective, the understanding of NTS it is not only individual, once a collaborative working environment is also necessary, considering the context of the whole sociotechnical system (Wachs et al., 2012). And, at the same time, workers shape their individual behaviour to be part of the team, characterizing this duality between individual and collective.

One important notion is that, according to the definition, NTS is not separate or in competition with technical skills. Rather, the two skill sets should be thought of as related and complementary. For instance, in healthcare domain, (McCulloch et al., 2009; Raison et al., 2018) indicates that NTS training can improve both technical and non-technical performance, increasing consequently patient safety. In the maritime industry, both accident reports and studies indicate a strong link between accidents and the lack of sufficient NTS (Fjeld et al., 2018). Further, the presence of skills such as communication and teamwork promotes high degree of cooperation, confidence, knowledge sharing, and crewmembers confidence in the onboard bridgework (Tavacioglu & Gokmen, 2020). In fact, the shipping industry is today under considerable changes through an increased complexity in technology, (Praetorius et al., 2020), requiring a differentiated onboard performance from seafarers.

2.2 Resource management training approaches

The simulation environment is a place where the work activities can be developed, in a training perspective, allowing the assessment and improvement of desirable skills, which are needed at the sharp end. Simulations are an effective educational tool used to acquire knowledge, skills and attitudes needed for performing jobs in different industries, but especially in the maritime domain (Håvold et al., 2015). Integrated with the training on actual equipment and training onboard ships and offshore facilities, the use of simulators helps to increase the training effectiveness and preparedness for real work scenarios (Mindykowski, 2017). Figure 1 presents an example of a bridge simulator.

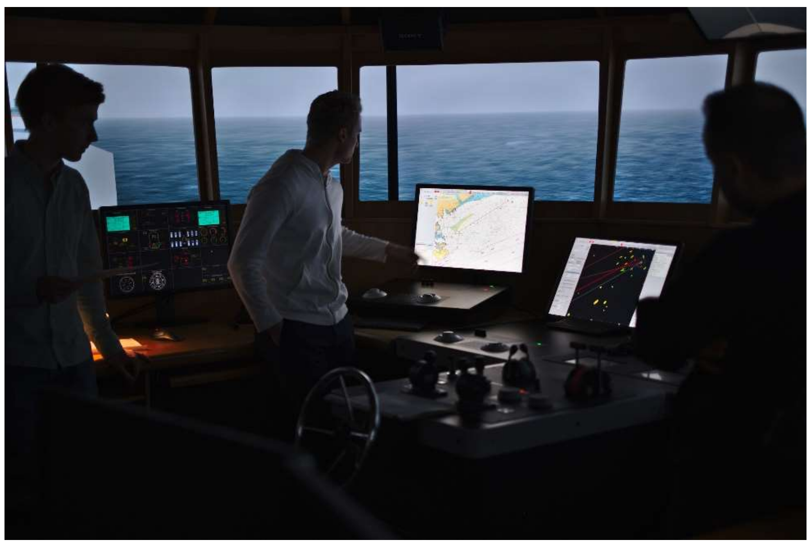


Figure 1. Example of a bridge simulator.

Using simulators in MET provides an opportunity to attain experience and to gain the necessary skills. Bridge Resource Management (BRM) has been developed as a maritime approach to Crew Resource Management (CRM). CRM has evolved from the 1970s onwards in the civil aviation industry, when it was recognized that many aircraft incidents were allegedly due to human error, rather than technical aspects of flying (Grech et al., 2008). CRM is a training precipitously designed to support the development of interpersonal aspects of flight operations, firstly focused on interactions within the cockpit, largely due to the Tenerife accident in 1977 (Alexander, 2015). NASA introduced its first CRM training as outcome of a human factors workshop in the beginning of the 1980s (Helmreich & Foushee, 2010). The maritime version of the CRM course was developed 1992, focused on bridge operations, similar with the first version of the CRM in aviation, focused in the cockpit (Hayward & Lowe, 2010). Soon later, IMO introduced the concept of BRM into the STCW Code, since IMO itself consider this training as the most adequate for onboard teams. The main reason for this is that particular training improves seafarers communication, teamwork, and their ability to use all available resources, in normal and degraded situations (Cavaleiro et al., 2020). Although, in the first instance, BRM was developed to enhance the relationship of the captain and pilot in the bridge, however, it evolved and became more embracing and focused on the safety and performance of all crew members (O'Connor, 2011).

After almost a decade of delivering BRM training in many locations around the world, a decision was taken to revise and expand the BRM training program. In 2003 the organizations involved in the global delivery of BRM training (The Swedish Club and BRM licensees) decided to rebrand the course from Bridge Resource Management (BRM) to Maritime Resource Management (MRM), to more accurately reflect the contents and objectives of the recently revised training program (Hayward & Lowe, 2010). Indeed, the latest revision of the STCW in 2010 increased the demands on NTS for officers onboard, which are normally trained as part of Bridge or Engine room Resource Management (STCW A- II/I, A-II/2, A-III/2, A-III/6), also called MRM (Praetorius et al., 2020). Since its origin, the MRM course was designed for deck and engineering officers, maritime pilots and shore-based personnel, having as main objective the development of skills needed to promote safety and responsible behaviour in everyday operations, minimising the risk of incidents onboard.

2.3 Briefing and debriefing in the maritime domain

In BRM, debriefing and feedback is an important and critical step in the evaluation of work performance and efficiency during exercises, simulations, and onboard training (Röttger et al., 2016). It is an after-action technique which brings the necessary elements for learning through doing. In specific cases, debriefing is considered to be the most critical part of the simulation experience (Crookall, 2013). Thus, a thorough debriefing of each simulation can allow the recognizing of uncovering instances where contextual, personal, and cultural elements influenced worker's performance (Fjeld & Tvedt, 2020). At the same time, it is also possible to identify the adjustments done by crew to achieve the desirable performance, which ultimately refers to the preparedness for outcomes and system resilience. Debriefing tools, as an AAR method, are important processes for knowledge gathering from what is really happening in work related activities, enabling the identification and management of interactions, skills, and abilities. After-action reviews are a principal means for an organization to assimilate difficult feedback about what hasn't worked (Kellen, 2008).

The debriefing is a post-experience analysis and team reflection of some activities or scenario, where it is possible to learn from the experience, allowing an integration between theoretical knowledge and practical experience onboard (Sellberg & Lundin, 2018). It is also noticed that debriefs are designed essentially to serve developmental purposes rather than evaluative or judgmental purposes, which not only yields more accurate feedback, but fosters circumstances that encourages knowledge exchange and perspective taking, maximizing experience learning from what really happened (Tannenbaum & Cerasoli, 2013). The learning fostered by debriefing is a part of the cycle of continuous improvement in the safety of work activities, providing and recognizing the necessary skills for this. As such, every crew member now had not only the right but also the responsibility of speaking whenever they considered necessary for safety, having their team leaders encouraging and rewarding this type of behaviour in their crew members (Ornato & Peberdy, 2014).

2.3 Resilience and safety in complex workplaces

Resilience Engineering (RE) defines safety as the ability to succeed under varying conditions, being important both understand desired and undesired outcomes (Hollnagel et al., 2011). The ability for these distinct outcomes characterizes the resilience and build the everyday performance. Therefore, RE focuses on understanding the entire work systems, including the workers and their interaction in teams, considering everyday performance and operational adaptability, emphasizing on how safety is achieved (Praetorius et al., 2020).

Resilience, therefore, concerns the ability to recognize and adapt to handle unanticipated occurrences, demanding a dynamic shift in the process, strategies, and coordination. In this sense, it requires a dynamic monitoring of the boundaries of a system's functioning, managing and adjusting the adaptative capacity in face of variation and challenges. Resilience is often described with the help of four abilities; the ability to respond, the ability to monitor, the ability to learn and the ability to anticipate (Hollnagel et al., 2011). These four cornerstones are essential for a system to be able to recognize challenging conditions, respond to them, evaluate the response, and prepare for future events. The four abilities are mutually dependent, and each represents one facet of a system's functioning (Praetorius et al., 2020).

3 MATERIALS AND METHODS

3.1 Resource settings

The debriefing tools has been developed within the settings of an advanced ship handling course for fourth year students in a maritime degree program at Kalmar Maritime Academy. Within this advanced course, BRM is taught through a mixture of classroom-based lectures and simulation exercises in a full-mission bridge simulator. The simulations comprise approximately 15 hours training in which different tools from the classroom-based education are meant to be trained in settings that is as realistic as possible. Nautical Science students and marine engineering students are hence trained together in teams.

Due to the current pandemic, however, the teams do not consist of five, but of two members: one future bridge officer and one future engineer. Each exercise ends with a debriefing. Normally the debriefing consists of two parts. The teams have their own debriefing in the bridge simulator, which is then followed up by an instructor-led debriefing with the whole class.

3.2 Debriefing tool

The debriefing tool (fig. 2) was developed in an iterative design approach comprising observations and contextual interviews with the simulator instructor to ensure the relevance of the developed questions for the student's exercises. Classroom-based lectures were also followed to gain a better understanding of the research settings. The tool consists of ten questions that aim to trigger reflection and enable the students to discuss their joint team performance during the exercise with the aim to identify potential improvements and lessons to share with their peers. The questions in the tool are inspired by three of the four resilience engineering cornerstones: respond, monitor and learn.

Figure 2. The debriefing tool and its questions.

Debriefing

Debriefing is carried out by the **entire team** related with the end of the trip.

- **General questions**
  - How has the voyage been from the bridge team's perspective?
  - How has the voyage been from the engine-room team's perspective?
  - How have we handled different parameters (environment, safety, economy)?
- **Assessment of the voyage**
  - What has gone well during the voyage? Why?
  - Has anything out of the ordinary happened?
    - If so what and why?
- **Potential for improvement**
  - Which aspects of our teamwork should we improve?
    - Communication
    - Decision making
    - Cooperation bridge/engine – roles and responsibilities
    - Anything else?
  - Has there been any indication for potential near misses that we should be aware of for next time?
  - What examples / experiences can we share with
    - Other crew members / The crew
    - The shipping company?

In the development stage, the best practices of the maritime area were employed, associated with the onboard experience of a multidisciplinary team, who drafted a preliminary version that served as the core material for all the other stages of development. The questions of the debriefing tool, presented in figure 2, were inserted in the training system program, so after the finalization of the simulation, the students can access online all the requirements from the training, including too, now, the debriefing tool questions.

3.3 Participants and procedure

About 25 students used the debriefing tool during their training, using it through the simulation system. Then, in another moment, five students who used the debriefing tool during their training, participated in the focus group.
After eight simulation exercises with student-led debriefing had been observed, the developed tool was tested. In a classroom-based lecture, the students received a ten minutes long presentation with information about the project, the aim of the tool. During the following simulation, the tool was handed out to students at the end of exercises. Two members of the research team followed the debriefing of the students and took notes on the student's use of the tool.
Additionally, contextual non-structured interviews were conducted with the instructors, aiming comprehend the learning dynamics of simulations, contextualized in the concepts adopted by the research. Some of these instructors also participated in the development and validation stages, which was constructive for the discussions fostered by the interview. The insights from this part were noted for further analysis.

3.4 Focus group evaluation

The focus group enables study the ways in which individuals collectively make sense of a phenomenon and build meanings around it. It reflects the process through which meaning is constructed in everyday life and to that extent can be regarded as more naturalistic (Bryman, 2016). In this sense, the debriefing tool utilization was evaluated with the help of a focus group with five students. The focus group took one hour. The following three questions were used to evaluate the tool:

Question 1: Based on your experience, what is the role of debriefing onboard?

Question 2: How has this Debriefing Tool affected your debriefing?

Question 3: If you got to design a Debriefing Tool, what items/questions would you have in it?

The question 1 aimed to perceive the role of the onboard debriefing from student's perspective. The question 2 sought to evidence how the Debriefing Tool was used to perform the debriefing. The question 3 aimed to collect student's opinion and suggestion regarding the tool.

The study of the focus group was developed in a dedicated room, where the students were led by two facilitators. Before the start of the activities, a Power Point© slides was presented, explaining the purpose of the focus group, the recording of the session, and explaining about their voluntary participation. At any time, they could give up their participation, without the need for justification. A consent form, with this information and the contact details of the research group, was provided to the students, and stored after their agreement and signature. All five students agreed to the terms and signed the document. Shortly thereafter, the recording started, and the discussion was mediated by the facilitators, having 10 minutes for question 1, 20 minutes for question 2 and 20 minutes for question 3. During this discussion, the facilitators also took several notes. Some final remarks were done after the end of the discussion in question 3, closing the activities with the acknowledgement of the student's participation.

4 Results

The discussion in question 1 arose points regarding onboard specific competencies and a blaming culture when reporting fails or near misses. Competences such as communication, teamwork and decision-making were placed as important for their performance in the work. They highlighted the importance of the debriefing for learning from near misses as much as exceptional performance, bringing some new perspectives from both. Also was discussed that not all personal see the same importance in the debriefing, which can be repetitive and have not enough time to be adequately done. At a certain point, it was argued that debriefing is very important for activities that involve high risks at sea, such as the transfer of dangerous or flammable products from and between tankers and other ships, allowing a better understanding of the risks and the anticipation of critical situations.

The question 2 started a discussion on how a debriefing tool can improve the debriefing itself. They argued that it is possible, because the tool helps to challenge co-workers and enables a guided discussion in the team, and perhaps a deep conversation regarding something important, although most of the time there will be not enough time for this. They affirmed that in real world scenarios, this debriefing tool will certainly improve worker's competences, however if having the appropriate time for that. Moreover, they discussed that is necessary to have a balance in the debriefing: not so repetitive, and not being rare, once that a debriefing performed using a debriefing tool will improve competences in a regular basis. Indeed, they also argued that a debriefing is necessary in all situations, not only when something bad, or too good, happens, because when not unusual happens, different perceptions give new perspectives for others, helping the crew to be ready for different work demands. Lastly, the debriefing should be performed in work hours, not out of the shift or in rest time.

In the discussion of question 3, they argued that a debriefing tool should be easy and simple to follow, like a small guideline, a short template, allowing a consistency feedback, helping to understand the work and communicate well. They pointed that a debriefing tool should have as less as possible question, being interesting, short and not too much, basically formed by three main sections: 1) What did good? 2) What did not so good? 3) How to improve from both of that? Also, once more, the discussion regarding the balance and time arouse, together with the onboard blaming culture. They were emphatic in saying that the blaming culture onboard do not help to have an effective debriefing, blocking the learning from mistakes and daily routine once the crew is afraid to report. To solve this, they pointed that maybe a "Debriefing box" could be useful, where crew members could deposit near misses and their mistakes without being exposed of blame and its consequences. In the end of the discussion, the students affirmed that a debriefing tool helps to do a better debriefing and evolve at work.

5 Discussion

The debriefing tool was performed by students based on their training in a maritime simulator centre at a university which has educational programs dedicated for MET. These particular students are part of the bridge and engine team officer's training program. The aim of this focus group was to encourage them to discuss about their debriefing and have insights of how this tool and performance can be continuously improved. Indeed, simulator trainings allow students to understand and analyse several work scenarios, improving performance in similar situations in the future (Wahl et al., 2020). This understanding of how the work happens, and the learning from this, was pointed out by the students in question 1 and 2 discussions, when they highlighted the importance of learning from near misses as much as exceptional performance. The active reflection during debriefing and that sharing of experiences are important learning process (Sellberg & Wiig, 2020), allowed by the utilization of this debriefing tool.

The learning promoted by the debriefing tool is the fundamental goal of simulator trainings, such as BRM. Also, learning is one of the resilience engineering cornerstones and is regarded as crucial for resilient performance (Wahl et al., 2020). In this sense, some studies place system resilience capacities are intrinsic related with activities or skills of sharpend operators, such as deck officers, which can be enhanced by planning or training (Bergström et al., 2015). In the question 2, the students pointed that a debriefing, applied after work activities independently of its nature, can allow different perceptions and discussions, helping the crew to be ready for different onboard demands. This ability to respond for different demands is another resilience engineering cornerstone, which allows the response for expected and non-expected changes, enhancing onboard crew performance by enabling their

experience learned (Hollnagel et al., 2011). In this context, resilience is achieved by understanding and managing the work variabilities that can be learn from debriefing, rather than by constraining it (Wahl et al., 2020).
However, the blaming culture when reporting fails and near misses, evidenced by the students in question 1 and 3, instead of enhance performance, can result in the opposite, constraining the learning and the ability to respond. In fact, the learning from errors in training simulation, without blaming, can allow positive outcomes for the entire class (Sellberg, 2018). As pointed by the students, this culture jeopardizes the debriefing, blocking the learning from mistakes and daily routine once there is a fear to report. In this sense, it is important to mention that under the resilience engineering concepts, the no-blame culture allows a systematic organizational learning, especially in the sharpend activities (Gattola et al., 2018). An onboard culture where the learning from the work comes from desired and undesired outcomes, not only helps to improve performance, but also build trust among crew members. Besides, other worker's competences were discussed in the focus group, stimulated by the debriefing.
In all discussions from question 1 to 3, some competences were mentioned as important for their performance in their onboard activities. In the question 1, communication, teamwork and decision-making arose as examples of these competences, being also known as NTS. The knowledge and application of NTS is currently inserted in the BRM training, where can be developed by simulations, debriefings etc (Hontvedt, 2015). The debriefing tool of this research is a way of development of these competences, pointed by the students as an adequate tool to assess the risks at sea, such as the transfer of dangerous or flammable products, allowing the anticipation of critical situations. The ability to anticipate is also another resilience engineering cornerstone, which enables a potential of preparation for future events, threats and opportunities (Hollnagel et al., 2011). The communication, discussed by the students in all three questions, is one of these NTS that can be continuously improved during the utilization of a debriefing tool, and should flow in a simple, clear and precise way (Fjeld & Tvedt, 2020).

6 CONCLUSIONS

This study analysed the utilization of a debriefing tool in simulator-based training classes and in a focus group, based on the resource management training approaches and the resilience engineering cornerstone concepts. Some individual competences, known as non-technical skills, are relevant for a safe onboard work and can be developed by a debriefing tool. There are some limitations in performance a debriefing, being the onboard blaming culture the responsible for the fear to report, blocking the learning from mistakes and daily routine. Despite the constrains, the debriefing tool was evidenced and ratified by the focus group as an adequate tool to learn and evolve in the work, acquiring knowledge and experience from work activities independently of its nature. This tool encourages a discussion regarding near misses and what happened of good, letting an open channel for feedbacks. Thus, it is possible to have a continuously learning from debriefing, allowing students to understand the work scenario, having insights, and acquiring experience which will improve their performance and prepare them for work situations in the future.

ACKNOWLEDGEMENTS

This article is part of the SjöResA-project financed by the Swedish Mercantile Marine Foundation. The authors would like to express their gratitude to the foundation for their support, as well as for the students who contributed to this research.

REFERENCES

Alexander, D. Y. (2015). *Never Wait for the Fire Truck*. CreateSpace Independent Publishing Platform.

Bergström, J., Van Winsen, R., & Henriqson, E. (2015). On the rationale of resilience in the domain of safety: A literature review. *Reliability Engineering and System Safety*, *141*, 131–141. https://doi.org/10.1016/j.ress.2015.03.008

Bryman, A. (2016). *Social Research Methods* (5th Ed.). Oxford University Press.

Cavaleiro, S. C., Gomes, C., & Lopes, M. P. (2020). Bridge Resource Management: Training for the Minimisation of Human Error in the Military Naval Context. *Journal of Navigation*, *73*(5), 1146–1158. https://doi.org/10.1017/S0373463320000235

Crookall, D. (2013). Climate Change and Simulation/Gaming: Learning for Survival. *Simulation and Gaming*, *44*(2–3), 195–228. https://doi.org/10.1177/1046878113497781

Fjeld, G. P., & Tvedt, S. D. (2020). How do BRM-training participants understand non-technical skills? *WMU Journal of Maritime Affairs*, *19*(2), 235–269. https://doi.org/10.1007/s13437-020-00198-9

Fjeld, G. P., Tvedt, S. D., & Oltedal, H. (2018). Bridge officers' non-technical skills: a literature review. *WMU Journal of Maritime Affairs*, *17*(4), 475–495. https://doi.org/10.1007/s13437-018-0158-z

Flin, R., O' Connor, P., & Crichton, M. (2016). *Safety at the Sharp End: A Guide to Non-Technical Skills*. Taylor & Francis Group, CRC Press.

Gattola, V., Patriarca, R., Tomasi, G., & Tronci, M. (2018). Functional resonance in industrial operations: A case study in a manufacturing plant. *IFAC-PapersOnLine*, *51*(11), 927–932. https://doi.org/10.1016/j.ifacol.2018.08.489

Grech, M. R., Horberry, T. J., & Koester, T. (2008). *Human Factors in the Maritime Domain* (1st Editio). Taylor

& Francis Group, CRC Press.
Håvold, J. I., Nistad, S., Skiri, A., & Odegård, A. (2015). The human factor and simulator training for offshore anchor handling operators. *Safety Science*, *75*, 136–145. https://doi.org/10.1016/j.ssci.2015.02.001
Hayward, B. J., & Lowe, A. R. (2010). The Migration of Crew Resource Management Training. In *Crew Resource Management* (2nd Ed., pp. 317–342). Academic Press. https://doi.org/10.1016/B978-0-12-374946-8.10012-3
Helmreich, R. L., & Foushee, H. C. (2010). Why CRM? Empirical and Theoretical Bases of Human Factors Training. In *Crew Resource Management* (2nd Ed., pp. 3–57). Academic Press. https://doi.org/10.1016/B978-0-12-374946-8.10001-9
Hollnagel, E., Pariès, J., Woods, D., & Wreathall, J. (2011). *Resilience Engineering in Practice: A Guidebok* (1st ed.). Ashgate Publishing Ltd.
Hontvedt, M. (2015). Professional vision in simulated environments - Examining professional maritime pilots' performance of work tasks in a full-mission ship simulator. *Learning, Culture and Social Interaction*, *7*, 71–84. https://doi.org/10.1016/j.lcsi.2015.07.003
IMO. (2010). STCW - A Guide for Seafarers - Taking into account the 2010 Manila amendments. In *International Maritime Organization*. STCW - Standards of Training, Certification and Watchkeeping for Seafarers.
Kellen, V. (2008). After-Action Reviews in IT. *Business-IT Strategies*, *11*(11), 12.
Le Coze, J.-C. (2021). *Post Normal Accident - Revisiting Perrow's Classic* (1 st). Taylor & Francis Group, CRC Press.
McCulloch, P., Mishra, A., Handa, A., Dale, T., Hirst, G., & Catchpole, K. (2009). The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Quality and Safety in Health Care*, *18*(2), 109–115. https://doi.org/10.1136/qshc.2008.032045
Mindykowski, J. (2017). Towards safety improvement: implementation and assessment of new standards of competence for Electro-Technical Officers on ships. *Maritime Policy and Management*, *44*(3), 336–357. https://doi.org/10.1080/03088839.2016.1275861
O'Connor, P. (2011). Assessing the effectiveness of bridge resource management training. *International Journal of Aviation Psychology*, *21*(4), 357–374. https://doi.org/10.1080/10508414.2011.606755
Ornato, J. P., & Peberdy, M. A. (2014). Applying lessons from commercial aviation safety and operations to resuscitation. *Resuscitation*, *85*(2), 173–176. https://doi.org/10.1016/j.resuscitation.2013.10.029
Praetorius, G., Hult, C., & Österman, C. (2020). Maritime resource management: Current training approaches and potential improvements. *TransNav*, *14*(3), 573–584. https://doi.org/10.12716/1001.14.03.08
Raison, N., Ahmed, K., Abe, T., Brunckhorst, O., Novara, G., Buffi, N., McIlhenny, C., van der Poel, H., van Hemelrijck, M., Gavazzi, A., & Dasgupta, P. (2018). Cognitive training for technical and non-technical skills in robotic surgery: a randomised controlled trial. *BJU International*, *122*(6), 1075–1081. https://doi.org/10.1111/bju.14376
Röttger, S., Vetter, S., & Kowalski, J. T. (2016). Effects of a classroom-based bridge resource management training on knowledge, attitudes, behaviour and performance of junior naval officers. *WMU Journal of Maritime Affairs*, *15*(1), 143–162. https://doi.org/10.1007/s13437-014-0073-x
Sellberg, C. (2018). From briefing, through scenario, to debriefing: the maritime instructor's work during simulator-based training. *Cognition, Technology and Work*, *20*(1), 49–62. https://doi.org/10.1007/s10111-017-0446-y
Sellberg, C., & Lundin, M. (2018). Tasks and instructions on the simulated bridge: Discourses of temporality in maritime training. *Discourse Studies*, *20*(2), 289–305. https://doi.org/10.1177/1461445617734956
Sellberg, C., & Wiig, A. C. (2020). Telling Stories from the Sea: Facilitating Professional Learning in Maritime Post-Simulation Debriefings. *Vocations and Learning*, *13*(3), 527–550. https://doi.org/10.1007/s12186-020-09250-4
Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors*, *55*(1), 231–245. https://doi.org/10.1177/0018720812448394
Tavacioglu, L., & Gokmen, N. (2020). *The non-technical skills of bridge officers on ship navigation : An assessment in terms of interpersonal skills*. *December*. https://doi.org/10.17485/IJST/v13i46.1646
Wachs, P., Righi, A. W., & Saurin, T. A. (2012). Identification of non-technical skills from the resilience engineering perspective: A case study of an electricity distributor. *Work*, *41*(SUPPL.1), 3069–3076. https://doi.org/10.3233/WOR-2012-0565-3069
Wahl, A., Kongsvik, T., & Antonsen, S. (2020). Balancing Safety I and Safety II: Learning to manage performance variability at sea using simulator-based training. *Reliability Engineering and System Safety*, *195*(October 2019), 106698. https://doi.org/10.1016/j.ress.2019.106698

Nine Principles for Supporting Distributed Sensemaking

Simon ATTFIELD[a], Shailey MINOCHA[a], Emmeline ELLIOTT[b], Bob FIELDS[a], Chris BABER[c], Rob HUTTON[d], Andrew LEGGATT[e], Peter COWELL-SMITH[f] and James HARRYMAN[g]

[a]Middlesex University London, UK
[b]Defence & Security Analysis Division, Dstl, UK
[c]University of Birmingham, UK
[d]Nottingham Trent University, UK
[e]Trimetis Ltd.
[f]MASS Ltd.
[g]British Army

ABSTRACT

The future military operating Environment (FOE) is likely to be one in which operational military units will be less dependent on traditional command and control (C2) structures, operating more as semi-autonomous, self-organising teams. The devolution of C2 functions places a greater emphasis on the need for local peer-to-peer sensemaking and decision making. The work in this paper is a result of a project looking at the idea of sensemaking as a distributed phenomenon in military settings. We discuss this changing context and, as a prelude to the principles, offer a characterisation of distributed sensemaking as a form of distributed cognition which involves the harnessing of disparate information resources. Based on a review of sensemaking and team sensemaking sources from areas including Organisational Studies, Computer Supported Collaborative Work, Naturalistic Decision Making and Human Computer Interaction, we propose nine principles for supporting distributed sensemaking based on specific identified challenges and potential interventions. The nine principles broadly relate to the provision of information, interpretation and communication.

KEYWORDS

sensemaking; distributed sensemaking; data-frame theory; command and control; future operating environment

INTRODUCTION

Sensemaking has been described as the motivated, continuous effort to understand connections between people, places, and events in order to anticipate their trajectories and act effectively (Klein, Moon & Hoffman, 2006). It has also been described as a quest for coherence between ideas at different levels of abstraction (Attfield, Fields & Baber, 2018). Sensemaking draws upon skills, knowledge and behaviours used to determine when a routine situation is moving to non-routine or when the situation may not be as originally perceived. Whilst much has been written about sensemaking, less has been said about sensemaking as a 'distributed' process (exceptions include: Fisher, Counts & Kittur, 2012; Attfield, Fields, Wheat, Hutton, Nixon, Leggatt & Blackford, 2015; Wheat, Attfield & Fields, 2016; and Attfield et al., 2018), although like many forms of cognition it is frequently a distributed phenomenon. As a form of cognition, sensemaking can be distributed in the sense that its representations and processes occur across the social interaction of individuals, the artefacts they use, the environments in which they work and previous events and experiences (Hutchins, 1995; Rogers, 2012; Hollan, Hutchins & Kirsh, 2000). Arguably, where sensemaking involves interaction between people and between people and artefacts, a full explanatory account could only be provided from such a point of view.

The work presented in this paper is a result of a research project funded by and conducted in collaboration with the UK Defence Science and Technology Laboratory (Dstl) which is looking at the idea of sensemaking as a distributed phenomenon in military settings. One of the aims of this research programme is to derive a set of principles that can be used to guide the design and assessment of distributed military sensemaking environments. Motivation for the project comes from the expectation that the future military operating environment will be one in which operational units are increasingly mobile and geographically dispersed with more decentralised command and control (C2) structures. It is also expected that units will be increasingly agile with greater peer-to-peer decision-making and sensemaking, in contrast to more traditional hierarchical C2 structures. This, combined with sensemaking challenges inherent within modern below-threshold and multi-domain conflict arenas, motivates the consideration of how distributed sensemaking can be facilitated in such contexts.

A review of sensemaking and team sensemaking sources was conducted with the aim of deriving a set of principles concerning supporting interventions applicable to this new context. The review considered sources from areas including: Organisational Studies, Computer Supported Collaborative Work, Naturalistic Decision Making and Human Computer Interaction. The scope extended beyond the military domain to include domains such as policing, legal investigations, and healthcare. The principles were derived from studies and models where substantive claims about sensemaking suggesting interventions or adaptations in areas such as tools, strategies and resources that might enhance or limit sensemaking. Hence the principles are framed as imperatives, with each based on a combination of findings or observations. The sources that we use either report naturalistic or *in vivo* studies, or models based on them. Naturalistic study is arguably the correct place to study a rich cognitive phenomenon such as sensemaking, or indeed, distributed sensemaking. This is not to say that experimental studies would not be possible or helpful, but it is the case that few exist. We return to this and the role that the principles may play in such an approach in the discussion.

We begin by elaborating the military context that provides the backdrop for the work. We then discuss sensemaking as a form of distributed cognition, including a characterisation of distributed sensemaking which draws on the idea that distributed cognition can be described as operating over a set of abstract information resources (Wright, Fields & Harrison, 2000; Attfield et al., 2018). We then describe the nine principles.

SENSEMAKING IN AN EVOLVING MILITARY CONTEXT

The future military operating environment is likely to be one in which operational combat units are required to be increasingly mobile and geographically dispersed with more decentralised C2 structures. One approach to the challenge of maintaining agility and resilience in the new context is for operational military units to be less dependent on C2 structures and to increasingly operate as semi-autonomous self-organising teams (the embodiment of "Mission Command" and a key tenet of the UK's STRIKE C2 concept (JCN 2/17)). The devolution of C2 functions, however, places a greater emphasis on the need for local and collaborative sensemaking capabilities to interpret information and assess potentially complex and unfolding situations so as to adapt and act effectively within the constraints of a broad military intent. C2 structures may need to change but sensemaking will be relevant to both the current and future structures. In fact, the STRIKE concept isn't just about a new physical paradigm, it also emphasises the need for a new mind-set: a new tactical mind-set to move, think and act faster and further than the enemy. It is this human mind-set that exploits speed, mobility and agile command to outpace the enemy even when information is denied or degraded and in a distributed sense. Our thinking on distributed sensemaking in this research programme directly addresses some of the cognitive aspects of the STRIKE concept.

If we see sensemaking as distributed and collaborative then we require approaches which can measure and change the ways in which military personnel gather and process information (training), the ways in which they understand information and how to interpret it (culture), the ways in which information is shared and interpreted (processes), and the ways in which information sharing is supported (technology). Distributed sensemaking is critical in support of coordinated actions by disparate units of action.

DATA-FRAME THEORY AND DISTRIBUTED SENSEMAKING

We begin by introducing some ideas and terminology that will be helpful for understanding the principles. These are Klein's Data-Frame model of sensemaking (Klein, Phillips, Rall & Peluso, 2007), 'Distributed Sensemaking Units', and 'Distributed Sensemaking'.

Data-frame theory of sensemaking

Klein's Data-frame theory of sensemaking (Klein et al., 2007) distinguishes two kinds of entity which interact during sensemaking: data and frame. Data are aspects of the world which a sensemaker experiences. A frame is a representation in the mind of the sensemaker which stands for the situation, for example, a doctor's beliefs about a patient's medical condition, a pilot's understanding of location and heading, or a warship captain's beliefs about the position, heading, and intent of an approaching aircraft. The frame acts as both interpretation and explanation of the data by accounting for it within a more integrated and complete picture. Importantly, the frame extends beyond the data, using background knowledge and expectations to fill gaps, or rather, it creates gaps in which data can be accommodated.

Klein et al. (2007) suggest that frames take a number of forms, including stories, explaining the chronological and causal relationships between events, maps describing locations and directions, or plans describing a sequence of actions. The term frame is intended as a synthesis of various concepts which have been used previously in accounts of human understanding, including frames by Goffman (1974), Minsky (1974), scripts by Schank & Abelson (1977), and schemata by Bartlett (1932), Neisser (1976), and Piaget (1952, 1954).

The Data-Frame theory presents sensemaking as a process of framing and re-framing in the light of data. As we encounter a situation, a few key elements or anchors invoke a plausible frame as an interpretation of the situation. Active exploration guided by the frame then elaborates it or challenges it by revealing inconsistent data. By extending further than the data, a frame offers an economy on the data required for understanding, but also sets up expectations. Hence a frame can "direct" information search and in doing so reveal further data that changes or shapes the frame. An activated frame acts as an information filter, not only determining what information is subsequently sought, but also affecting what aspects of a situation will subsequently be noticed.

Important to the selection of a frame is the sensemaker's repertoire of frames and this underlies a distinction between experts and novices. It is argued that whilst experts and novices reason using the same procedures, experts have richer repertoires of frames that are better differentiated, allowing sense of a greater variety of situations, to be more precise about the situations, and to focus on fewer (but higher level) elements in a given situation (Lipshitz & Shaul, 1997). Klein et al., (2007) also argue that frame activation depends upon the sensemaker's 'stance' including factors such as workload, motivation, and their current goals.

Characterising Distributed Sensemaking

We use the term 'Distributed Sensemaking Unit' (DSM Unit) to refer to an individual, a semi-autonomous or autonomous system, who can perform a sensemaking function. We also see a network of DSM Units behaving as a single system, and with purpose as a recursively defined unit. Following Attfield et al. (2018), we associate a DSM Unit with a set of abstract information resources which enable and motivate ongoing sensemaking processes (see Figure 1). These information resources include: (1) Data (that is, information about the world); (2) Expertise and Domain Knowledge (including frame repertoires, used for making sense of data); (3) Situation Picture (a model of the current situation derived, that is, inferred from 1 and 2); and, (4) Values and goals (high-level interests that define the kind of sense that might be of value and through this can be translated into goals, that is, things that can be achieved).

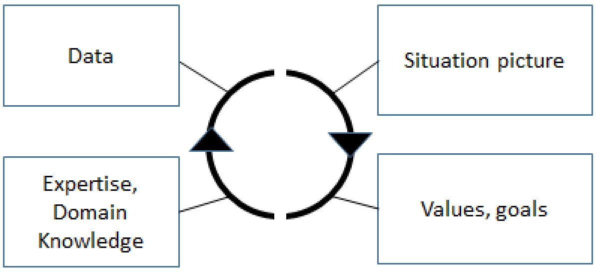


Figure 1. A Distributed Sensemaking (DSM) Unit

In sensemaking, data, expertise/domain knowledge, situation picture and value/goals can be thought of as abstract information resources that are harnessed in the service of sensemaking (Attfield et al., 2018). More specifically: new *data* can trigger uncertainty concerning a *situation picture*. This might relate to the past, present or future state of physical things, or the actions/intentions of agents, or all of these. This can lead to uncertainty in decision making, given a set of *values,* meaning it can become unclear how to operationalise these values as goals (that is, things that can be done). Reasoning involves the interpretation of the situation given *data* and *expertise/domain knowledge*. Multiple interpretations or *situation pictures* might be considered, each of which may give priority or credibility to different *data* and/or *expertise/domain knowledge*. The aim is to reach a view where *data*, *expertise/domain knowledge* and *situation picture* form a coherent whole and support purposeful action - which may involve affecting the situation, seeking further information, or doing nothing.
An account of sensemaking in terms of resources such as this provides a basis for a characterisation of distributed sensemaking in the sense that where sensemaking is distributed, *data*, *expertise/domain knowledge*, *situation picture*, and *values/goals* may be distributed across multiple parties (or DSM units) or artefacts. For example, one party may receive new *data*, whilst others may be a source of relevant past *experience/knowledge.* Parties may form more or less local or global *situation pictures* and different parties may make more or less local or global decisions. Viewed in this way, a challenge for DSM is to harness disparate resources in the interests of collective meaning making and purposeful action, placing an emphasis on communication and the negotiation of meaning.

Deriving the Principles

A literature review was conducted with the scoping goal, 'to identify potential barriers and facilitators for distributed sensemaking'. For the most part, searching was conducted using Middlesex University's federated search engine. This searches online databases, books and theses according to the University library's subscriptions and holdings, including academic databases in Computer Science (e.g. ACM DL, IEEE Xplore) and Psychology (e.g. APA PsycArticles, Science Direct, Web of Science).
The following search terms were used individually and in combination:

- Keywords: 'situation awareness', 'sensemaking, 'situation assessment', 'team situation awareness', 'distributed cognition', 'teamwork', 'data-frame theory of sensemaking', 'problem detection sensemaking', 'command and control sensemaking', 'sensemaking decision-making', 'overt sensemaking', 'tacit sensemaking', 'crisis management', 'distributed sensemaking military', 'collaborative sensemaking', 'common ground', 'team-learning command and control', 'team sensemaking', 'distributed collaborative analysis', 'geographically distributed sensemaking', 'naturalistic decision-making', 'organised sensemaking', 'team resilience', 'macrocognition', 'team adaptation', 'adaptive expertise'.

Resulting papers were ranked by relevance by the search tool. This was used to bulk eliminate irrelevant documents. This was done by examining the titles and abstracts of each paper in turn, starting from the most highly ranked paper, judging for relevance against the scoping objective, and applying a stopping rule based on a judgement of diminishing returns (usually around the 30th-50th document). Documents judged to be relevant were downloaded. A further triage step was then conducted based on a reading of introductions, methods and conclusions. The final set of papers were exported to a bibliographic reference management system.

In addition, ad hoc searching was conducted using Google and Google Scholar using similar keywords, and also new documents were obtained through reference chaining and focussing on selected authors known for their work in the area. In some cases, this involved searching for known authors and ideas, drawing on expertise within the research team. Finally, the process was occasionally repeated for emerging sub-themes.

Literature sources were reviewed and emergent themes created based around the idea of 'barriers and facilitators for distributed sensemaking' (e.g. Klein, Pliske, Crandall & Woods, 2005; Weick, 2009). This led to the identification of phenomena relating to factors such as tools, strategies and resources that might enhance or limit sensemaking, with a focus on collaborative settings. This literature review alongside the authors' previous research in this area (e.g. Attfield et al., 2015; Attfield et al., 2018) guided the derivation of the principles which are elaborated in the next section.

Nine Principles for Supporting Distributed Sensemaking

In summary, the principles are (in no particular order):

1. Provide sufficient cues for sufficient sensemaking;
2. Support low cost information workflows;
3. Represent information quality and provenance;
4. Promote expertise/domain knowledge;
5. Allow time to acquire data/information to build an evidence-based and coordinated situation picture;
6. Use strategies for the negotiation of sense;
7. Where appropriate, use strategies for frame enumeration and elimination;
8. Provide explanatory context for actions, orders and requests; and,
9. Minimise the costs of achieving and maintaining common ground.

In the following, we elaborate the principles in turn.

1. Provide sufficient cues for sufficient sensemaking

The dynamic of sensemaking is to alternate between seeking and processing data and developing a frame where the frame offers a plausible explanation of the data (Klein et al., 2007). It has been described as the enlargement of small cues. The cue that alerts a DSM unit to a situation, however, is rarely definitive and can lead to additional information seeking to test uncertain frames or to disambiguate competing frames (e.g., diagnosis) (Attfield Pagnacco & Benedyk, 2011). Further, the range of critical situations can be relatively small, but it can be important to disambiguate them quickly to inform action. Depending on their roles, values and goals, DSM units operate more efficiently and more effectively through the prior anticipation of cues or cue-patterns relating to significant phenomena and using this to inform the training of personnel and artificial intelligence systems and capabilities for sensing, recording, and displaying data.

2. Support low cost information workflows

Sensemaking in information rich environments frequently involves the construction of information interaction workflows which involve, often repeated, sequences of manual, semi-automated or automated steps such as search, triage, extraction and structuring. Such workflows have been observed and described in intelligence analysis (Pirolli & Card, 2005), the construction of courseware (Russell, Stefik, Pirolli & Card, 1993), legal eDiscovery (Attfield & Blandford, 2011) and crime investigation (Selvaraj, Attfield, Passmore & Wong, 2016). The resulting information representations (for example, timelines, geospatial layouts, social networks, and command structures) can support insight, but conduct of the workflows can be labour intensive, imposing costs of time and physical and cognitive workload. It is important to anticipate workflow steps to provide artefacts and technologies that minimise the time and effort required.

3. Represent information quality and provenance

Intensive sensemaking can involve analysing and drawing insights from large amounts of information. Interpreting an assessment arising from such analysis typically requires knowledge of data sources and data quality, as well an understanding of the information transformation and inference steps performed during the analysis (Hutton, Attfield, Wiggins, McGuinness & Wong, 2011). Such histories have been referred to as 'data provenance' and 'analytic provenance' (sometimes 'insight provenance') respectively (Xu, Attfield, Jankun-Kelly, Wheat, Nguyen & Selvaraj, 2015). Data provenance information is a valuable resource for appreciating the identity and reputation of data sources and knowing how, when and where some data were acquired, and so forth. Analytic provenance is an important tool for both analysts and consumers who wish to reflect on issues such as how data quality and uncertainty propagate through an analysis as well as what has and has not been done in order to understand the meaning of the result. Analytic provenance information can also support audit and training analytic expertise after the fact.

4. Promote expertise/domain knowledge

Domain expertise is key to sensemaking. Domain experts reason in much the same way as novices, but training and experience provides domain experts with a wider repertoire of differentiated frames to support better sensemaking (Klein et al., 2007). They acquire a library of generic 'template' frames or belief structures about how the world tends to work in a given domain (Attfield et al., 2018). We suggest to that is necessary to make domain knowledge and expertise available in the minds of personnel, external artefacts and artificial intelligence support. Promoting domain knowledge can involve training on generic concepts relevant to a domain, for example, communicated and remembered through storytelling (Weick, 1995; Rantatalo & Karp, 2018) as a tool for 'sensegiving' (Gioia & Chittipeddi, 1991). In addition, domain knowledge does not reside purely in the head but can be embedded in cognitive tools such as artefacts that convert mental reasoning tasks into perceptual tasks or physical procedures (see for example, Wheat et al., 2016), or embedded in artificial intelligence tools used for recognition. Examples of the former include maps and data visualisations. Examples of the latter include image recognition systems.

5. Allow time to acquire data/information to build an evidence-based and coordinated situation picture

In high-stake situations when unexpected things happen, the pressure to respond can be high, despite possible conditions of information overload, ambiguity, confusion, and feelings of disorientation. Sensemaking often occurs in non-routine situations. In these situations, it can be helpful to temporarily withdraw, either metaphorically or physically, and pause - to give an opportunity for 'system 2 thinking' (Kahneman, 2012). Information should be requested and allowed to propagate across the distributed sensemaking unit in order to develop an evidence-based and coordinated situation picture, leading to coordinated action.

6. Use strategies for the negotiation of sense

Since background knowledge and experience is a significant factor in the range and differentiation of interpretative frames (Klein, 2006), alternative perspectives can offer an advantage in sensemaking. Overly hierarchical teams can lead to the deferral of sensemaking by junior members and the failure to declare critical information. Developing strategies for taking soundings from team members and also those outside the team who may offer a fresh perspective. Strategies have been described for accessing and managing alternative interpretations, such as the appointment of a Devil's Advocate (Klein et al., 2007; Klein, Wiggins & Dominguez, 2010; Hammond, Keeney & Raiffa, 2006), Alternatively Thinking Team (Houghton & Amey, 2018) setting rules for ritualised dissent (Klein et al., 2010), setting criteria for alarm raising, a system of rewards for encouraging and reporting of errors (Rochlin, 1996) and make the most of any failure that is reported (Weick, Sutcliffe & Obstfeld, 1999), review meetings and informal huddles (Attfield & Blandford, 2011). Different strategies may suit different situations.

7. Where appropriate, use strategies for frame enumeration and elimination

Sensemaking frequently involves abductive reasoning (reasoning to the best possible explanation) (Klein et al., 2007). However, abductive reasoning is 'defeasible', meaning its conclusions are not guaranteed, even if they may appear compelling. A frequent limiting factor is the small number of possibilities or 'frames' considered at a given time. Klein et al. (2007) speculate an upper limit of three. Confirmation bias involves giving serious consideration to only one. Where appropriate, use strategies for frame enumeration and elimination. Feltovich, Johnson, Moller & Swanson (1984) refers to such a set as a 'logical competitor set' and uses its extent as a measure of expertise. Analysis of Competing Hypotheses (ACH) is an elimination strategy proposed by Heuer (1999) intended to mitigate against confirmation bias. Signal's intelligence analysts have been observed to use elimination when drawing inferences from signals (Attfield et al., 2015). Such techniques frequently require the use of specially designed artifacts or technologies that lay out a logical competitor set against a set of features and allow the recording of elimination decisions.

8. Provide explanatory context for actions, orders and requests

Sensemaking is an iterative process of deriving meaning, acting, receiving information, deriving new meaning, and so on. Action is predicted on the meaning that is developed, but imperatives (actions, orders and requests for information) can lack context if the motivating sense or intent is missing. This context enables a recipient to make better sense of what is being asked and more adaptive and coherent response in the face of unanticipated factors.

Klein et al. (2010) notes that orders lacking rationale place a burden of sensemaking on subordinates who have to make adjustments without fully appreciating how the situation has changed. Where imperatives are passed down a chain of command, it is vital to provide the 'sense' that underpins them, either in terms of a motivating situation picture, or a broader plan, or objective.

9. Minimise the costs of achieving and maintaining common ground

For any two people, common ground is the sum total of the mutual knowledge, beliefs and assumptions that they share (Clark, 1996) and know that they share. Language is a form of collaborative action which is used to increase common ground, and on which coordinated action depends (Monk, 2009). Grounding takes effort, and communication technologies can increase or reduce the associated costs. It is important to aim reducing these costs through the use of standardised terminology, protocols and procedures, mission briefings, physical positioning for common visual and auditory context where appropriate. In the case of technologically mediated communication, design for shareable and coordinated information displays with tools for supporting pointing/indication of objects in the word or on a display, physically or verbally (known as 'deixis') are some of the strategies to reduce costs.

DISCUSSION

Sensemaking is a process of deriving meaning from a situation. It interacts iteratively with data driving a frame (or frames), and frame (or frames) driving the need for data – and this is in the context of some guiding interests and goals. The idea of a DSM Unit is used to explore and describe the minimum requirements of a sensemaking agent in terms of abstract information resources (which can be realised in the head or in the environment). It is also used to explain how actions of a collection of DSM Units with coordinated purpose and action (e.g. a commander's intent), can be thought of as DSM in itself. This idea simultaneously draws on both the idea of recursion in Stafford Beer's (1981) Viable Systems Model, and Dennett's (1987) Intentional Stance. This coordination of course requires either explicit or implicit communications, particularly in the face of unexpected conditions, which are after all, key drivers of sensemaking. In a network of such units, we anticipate both local and global sense emerging with coordination arising and then disappearing, as required, depending on local and global circumstances.

In broad terms, the principles that we have described derive from extant models and observational findings relating to the acquisition, enrichment, and manipulation of information, and strategies for interpretation, and communication. For the sake of coverage, they are intentionally high-level. They have also been identified with specific interest in relevance for military operating environment. However, given that they were derived from a variegated literature, we are optimistic about their general applicability.

Such work may be approached, for example, by translating the principles into guidelines or even interventions specific to fairly well characterised contexts. The principles could be used to guide evaluations of interventions technological or otherwise, or to analyse critical incidents or failures in distributed sensemaking, or to determine the success factors in success stories of distributed sensemaking. Understanding how the principles translate into specific contexts, however, is a further step that we have not addressed here. We expect that relevance may vary from domain to domain - for example, not all sensemaking domains have physical workflows in which information is manipulated in systematic, repeatable ways, although such workflows have been described as part of intelligence analysis (Pirolli & Card, 2005).

Returning to the question of the sources for the principles being naturalistic studies or models based on them, we believe that this affords the principles important ecological validity - an argument that will be familiar to the naturalistic decision making community. But there is also perhaps a question of the potential for experimental validation and the contribution that this might make to the epistemological status of the principles. We note here that, in general, little work has been done in moving sensemaking research beyond qualitative, observational and interpretive paradigms, and towards quantification and experiments, although this possibility is something that we see as making a potentially valuable contribution.

In terms of the principles, these can be framed as hypotheses that can be tested experimentally and in this sense provide a research agenda of sorts. We see this as a worthy move within a pluralistic research strategy, but one that is not without significant challenge. For example, an obvious challenge is to represent a natural setting with sufficient fidelity so that findings can be meaningfully extrapolated back into natural settings. Another challenge, particularly for studies involving group activity such as distributed sensemaking, is the practical problem of conducting experiments that are sufficiently rich and ecologically realistic whilst still providing enough data to draw generalisable conclusions. As a step in making inroads into this space, we are exploring approaches to experimentally operationalising selected principles in order to measure effects in the context of map-based military scenarios in a dispersed operational context in which participants will be wrestling with sensemaking challenges that typically occur in military environments.

ACKNOWLEDGMENTS

This work was funded by the Defence Science and Technology Laboratory (Dstl), UK Ministry of Defence.

REFERENCES

Attfield, S., Pagnacco, F. & Benedyk, R. (2011). When enough is enough: time-restricted sensemaking in underground railway control. In: *10th International Conference on Naturalistic Decision Making*, 31 May - 3 Jun 2011, Orlando, FL, USA.

Attfield, S. & Blandford, A. (2011). Making Sense of Digital Footprints in Team-Based Legal Investigations: The Acquisition of Focus, *Human-Computer Interaction: Sensemaking, 26 (1/2),* 38-71.

Attfield, S., Fields, B., Wheat, A., Hutton, R., Nixon, J., Leggatt, A. & Blackford, H. (2015). Distributed sensemaking: a case study of military analysis, *International Conference on Naturalistic Decision Making*, McLean, VA.

Attfield, S., Fields, B. & Baber, C. (2018). A resources model for distributed sensemaking, *Cognition, Technology & Work*, *20(4)*, 651-664.

Bartlett, F.C. (1932). *Remembering: a study of experimental and social psychology*, Cambridge: Cambridge University Press.

Beer, S. (1981). *Brain of the Firm*. Wiley.

Clark, H. (1996) *Using Language*, Cambridge University Press.

Dennett, D. C., (1987) *Three Kinds of Intentional Psychology*, pp. 43–68 in Dennett, D. C., *The Intentional Stance*, MIT Press.

Feltovich, P. J., Johnson, P. E., Moller, J. H. & Swanson, D. B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. *Readings in medical artificial intelligence*, 275-319.

Fisher, K., Counts, S., & Kittur, A. (2012). Distributed sensemaking: improving sensemaking by leveraging the efforts of previous users, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 247-256.

Gioia, D. A. & Chittipeddi, K. (1991). Sensemaking and sense giving in strategic change initiation. *Strategic Management Journal*, vol. 12, 433-448.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.

Hammond, J., Keeney, R. & Raiffa, H. (2006). The hidden traps in decision making', *Harvard Business Review*, *84(1),* 118-126.

Heuer, R. (1999). *Psychology of Intelligence Analysis. Central Intelligence Agency*, Center for the Study of Intelligence, Washington, D.C.

Hollan, J.; Hutchins, E. & Kirsh, D. (2000). Distributed Cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, *7*, 174–196.

Houghton, P. & Amey, D. (2018). Alternative Thinking Team. *The proceedings of the 23rd International Command and Control Research Technology Symposium, Multi-Domain C2.* Held at Florida Institute for Human & Machine Cognition (IHMC), Pensacola, Florida, USA, 6-9 November 2018.

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

Hutton, R., Attfield, S., Wiggins, S., McGuinness, B., and Wong, W. (2011). Team Sensemaking Assessment Method (TSAM): A human factors inspection method for evaluating collaborative analysis technologies. *TES Report 108183* (Deliverable under MoD CDE Contract 19998). BAE Systems, Advanced Technology Centre.

JCN 2/17, *Future of command and control* (2017). [Online] Available at https://www.gov.uk/government/publications/future-of-command-and-control-jcn-217 (Accessed 17 May 2021).

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin.

Klein, G., Pliske, R., Crandall, B. & Woods, D. (2005). Problem Detection. *Cognition, Technology & Work*, *7(1)*, 114-128.

Klein, G. (2006). The strengths and limitations of teams for detecting problems, *Cognition, Technology & Work*, *8(4),* 227-236.

Klein, G., Moon, B. & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, *21(4)*, 70-73.

Klein, G., Phillips, J.K., Rall, E.L. & Peluso, D.A. (2007). A data-frame theory of sensemaking. In Expertise out of context. In *Proceedings of the sixth international conference on naturalistic decision making*, 113-155. New York, NY, USA: Lawrence Erlbaum.

Klein, G., Wiggins, S. & Dominguez, C.O. (2010). Team sensemaking, *Theoretical issues in ergonomics science*, *11(4)*, 304-320.

Lipshitz, R. & Shaul, O. B. (1997). Schemata and mental models in recognition-primed decision making. *Naturalistic decision making*, 293-303.

Minsky, M. (1974). *A framework for representing knowledge*. MIT-AI Laboratory Memo 306.

Monk, A. (2009) *Common Ground in Electronically Mediated Conversation*. 1st edn. San Rafael: Morgan & Claypool Publishers.

Neisser, U., (1976). *Cognition and Reality*, San Francisco, CA: W.H. Freeman.

Piaget, J., (1952). *The origins of intelligence in children*. (M. Cook, Trans.). W W Norton & Co.

Piaget, J. (1954). *The construction of reality in the child*. (M. Cook, Trans.). Basic Books.

Pirolli, P. & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis*, vol. 5.

Rantatalo, O. & Karp, S. (2018). Stories of policing: The role of storytelling in police students' sensemaking of early work-based experiences. *Vocations and Learning*, *11(1)*, 161-177.

Rochlin, G. I. (1996) Reliable organizations: present research and future organisations. *Journal of Contingencies and Crisis Management*, *4(2)*, 55-59.

Rogers, Y. (2012). *HCI theory: Classical, modern, and contemporary*. Synthesis Lectures on Human-Centered Informatics. *5*, 1-129.

Russell, D. M., Stefik, M. J., Pirolli, P. & Card, S. K. (1993). The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, 269-276.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum.

Selvaraj, N., Attfield, S., Passmore, P. & Wong, B. W. (2016). How analysts think: think-steps as a tool for structuring sensemaking in criminal intelligence analysis. In *2016 European Intelligence and Security Informatics Conference (EISIC)*, 68-75.

Weick K. E. (1995) *Sensemaking in organizations*, Thousand Oaks, CA: Sage Publications.

Weick, K. E., Sutcliffe, K. M. & Obstfeld, D. (1999). Organizing for High Reliability: Processes of Collective Mindfulness. *Research in Organizational Behavior*, vol. 21, 81-123.

Weick, K. E. (2009) Managing the Unexpected: Complexity as Distributed Sensemaking, in *Making Sense of the Organization, in volume 2: The Impermanent Organization*, John Wiley & Sons, Incorporated.

Wheat A., Attfield S. & Fields B. (2016) Developing a model of distributed sensemaking: a case study of military analysis. *Informatics*, *3(1)*, 1-16.

Wright, P. C., Fields, R. E. & Harrison, M. D. (2000). Analyzing human-computer interaction as distributed cognition: the resources model. *Human-Computer Interaction, 15(1)*, 1-41.

Xu, K., Attfield, S., Jankun-Kelly, T. J., Wheat, A., Nguyen, P. H. & Selvaraj, N. (2015). 'Analytic provenance for sensemaking: A research agenda'. *IEEE Computer Graphics and Applications*, *35(3),* 56-64.

DSTL/CP130768

Adaptation in Team Performance for Sensemaking: Are We Ready For The Challenge ?

Alison CLERICI[a], Claire HILLYER[a], Leonie WEBSTER[b], Desmond McEWAN[c], Olivia BROWN[c] and Julie **GORE**[c]

[a]*QinetiQ, UK,*
[b]*Dstl, UK*
[c]*University of Bath, UK.*

ABSTRACT
This paper reports an on-going project to explore the challenge of measuring adaptive performance in teams for improved sensemaking. The paper describes an initial rapid review of literature which highlights: (i) the conceptual emphasis of the existing literature, (ii) nascent measures for examining individual adaptive skill and adaptivity in teams, and (iii) a lack of empirical evidence for the development or efficacy of training interventions to improve or develop adaptive team performance. We build on previous understanding about adaptive performance. Nascent measures of individual and team adaptivity are identified and evaluated. Initial experiences of efforts to experimentally measure adaptive team performance for sensemaking will be presented if available.

Keywords
Measurement, Adaptive Performance; Teams, Sensemaking

Contact Information of Author: j.gore@bath.ac.uk

INTRODUCTION & BACKGROUND

The concept of measuring adaptive performance at the individual level has been examined by cognitive psychologists, the military, and organizational/industrial psychology researchers in health, sport and high reliability contexts, to name a few for the past two decades. This work has had some success, and work continues to examine individual adaptive cognition at pace, with continued attempts to improve training to improve adapative performance (see Ward, Gore et al, 2017 for a review). Attempts however, to evaluate, measure, assess and train as a result of examing the current evidence base *of* and *for* adaptive team performance is still, despite the rhetoric which suggests otherwise, - a relatively nascent area.

Previous concerns have been expressed that the concept of 'adaptivity' is not an empirically robust one (Hutton et al, 2017, Ward et al, 2017) and we concur with this suggestion, alongside recognising the need to navigate a complex a messy and unstructured evidence base. What we do know is that there is signficant work which has examined behavioral markers for teams that work (see Salas et al 2007 to date; Flin, 2008); measuring adaptivity here however, suggest further layers of complexity.

The aim of this paper is to briefly review current evidence, highlighting and evaluating the potential strtenghts of developing measures of team adaptivity, with the aim of informing training to improve adaptive team performance for improved sensemaking for military operational context. Our key question here, is based on this evidence of reviewing academic and practitioner literature of adaptive team performance, and gaining the views of stakeholders – are we ready for the challenge ?

METHOD

We conducted a rapid scoping review of existing literature that has attempted to explore ways to measure and train adaptive team performance. We conducted searches using search-term combinations of library databases, and conducted an initial review of 1104 abstracts. From these abstracts, we subjectively identified those sources most relevant to our goals (*n* = 34). Based on subsequent reading of these and related articles, we collated a database of approximately 170 related publications. In addition we consulted with QinetiQ, Dstl, international academics and stakeholders working, researching and educating teams, in order to obtain diverse perspectives on adaptive team performance and training.

Two key findings derive from this review. First, the literature base on adaptive expertise continues to be largely conceptual. Empirical data are sparse at the individual level and almost non-existent at the level of the team. Second, adaptive performance training for teams has multiple interpretations and interpretations of "adaptivity"are nuanced in a number of ways. For example, we often need to establish if researchers are suggesting that training techniques are adaptive or that the behaviour and or cognition of adaptive teams is being examined via training. Either way we argue that reflections in both areas is required.

Research synthesis

Over 1000 resources were reviewed, from both the academic (journals and peer-reviewed journals and the applied (e.g. government technical reports) literature bases. A wide range of perspectives, disciplines and objectives were captured in the review, including work from learning, experimental, occupational (industrial-organisational) and human factors psychology, education, human resource mangement and training primarily but not exclusively.
The concept of adaptive team performance spans at least three key areas of literature including leadership and management, team behaviour and within different branches of psychology. Below in Table 1 we highlight some of the adaption/adaptivty/adaptive peformance definitions we identifed.

Table 1: Adaptation/Adaptivity/Adaptive performance definitions

Concept	Definition	Source
Adaptation	*"the ability to change work processes and the ability to change the organization"*	Alberts and Hayes (2003, p.128)
	"the ability to alter force organization and work processes when necessary as the situation and/or environment changes"	Alberts and Hayes (2003, p.152)
Adaptivity	*"the ability to employ multiple ways to succeed and the capacity to move seamlessly among them"*	Hoffman, et al. (2014, p.14)
	The capacity of a system to achieve its goals despite the emergence of circumstances that perturb the system toward the boundaries of its competence envelope. The work system is able to employ multiple ways to recover,	Hoffman, Hancock and Woods (2015)

Concept	Definition	Source
	or develop new ways to succeed, and can move seamlessly between them. The work system can reallocate and re-direct its resources and activities, retrench from the boundaries of its competence envelope, and achieve its goals.	
Adaptive performance	*"Timely changes in understanding, plans, goals, and methods in response to either an altered situation or updated assessment about the ability to meet new demands that permit successful efforts to achieve intent."*	Ward et al. (2016), p.3
	"cognitive, affective, motivational, behavioural modifications made in response to the demands of a new or changing environment, or situational demands"	Baard et al. (2014), p.50

Whilst definitions of adaptive performance across practitioner, military and academic literature vary in their emphasis the research team for this paper developed the definition below:

"Team adaptation involves the adjustments to actions that teams implement (team adaptation processes) as well as the result of those actions (adaptive team performance) in response to the demands of new, unforeseen, uncertain, or changing situations, environmental demands, or task barriers". (Clerici, Hillyer, McEwan & Gore, 2021, p.8)

ANALYSIS

Table 9 lists the self report measures we identified as potentially mapping with our definition. (Note behavioural measures are examined elsewhere and not in this paper).

We identified that the questionnaire employed by Rousseau & Aubé (2020) appears to be one of most appropriate self-report measure of team adaptation. Importantly, the development of this measure was guided by the conceptual framework of adaptation put forth by the important early work of Pulakos et al. (2000) and displays evidence of good reliability (Cronbach's alpha in that study was .84) and various aspects of validity (e.g. data from this measure were significantly related to data from measures of leader behaviours and shared team leadership; therefore, there is evidence of construct validity).

This measure consists of 7 items, each of which are rated on a 7-point response scale (1 = strongly disagree, 7 = strongly agree). The questionnaire taps into all dimensions of adaptability except 'demonstrating physically-oriented adaptability' which the authors deemed to be irrelevant to the study population. A sample item is *"Members of this team take effective action to deal with crises or urgent work demands"*. The measure can be completed by team supervisors/leaders (as was the case in the Rousseau & Aubé (2020) study) or by the team members themselves (which has been done with other measures in several other studies of adaptation, e.g., Abrantes et al., 2018; Georganta et al., 2020; Marques-Quinteiro et al., 2013).

Table 2: Self-report measures

Source	Decision	Justification for inclusion/exclusion
Self-report measures		
Sverdup et al. (2017)	No	Includes three general items. Does not tap into the various dimensions of team adaptation as specified by Pulakos et al. (2000) as the Rousseau & Aubé (2020) measure does. Tangential consistency with D11 working definition of team adaptation as it concerns how participants handle disruptions and their perceived adaptation within teams (not the adaptation of the teams themselves). Example item: *"All in all, I think the team handled the disruptive event in a good manner"*

Source	Decision	Justification for inclusion/exclusion
Konradt et al. (2015)	No	Includes three items. Not consistent with D11 working definition as the measure concerns the implementation of an agreed strategy rather than adapting to unforeseen changes. Example item: *"After agreements have been made in this team, everyone does things in the same manner"*
Wiedow & Konradt (2011)	No	Same measure as Konradt et al. (2015). Includes three items on team adaptation. Measures the implementation of an agreed strategy rather than adapting to unforeseen changes. Example item: *"After agreements have been made in this team, everyone does things in the same manner"*
Marques-Quinteiro et al. (2013)	No	Includes nine items. Consistent with the D11 working definition and includes three dimensions from the Pulakos et al. (2000) conceptual framework of adaptation. The Rousseau & Aubé (2020) was assessed as being more suitable as it considers seven of the eight dimensions from Pulakos et al. (2000) instead of three. Example item: Participants asked to rate how their teams solve problems creatively (*"using new ideas to overcome any challenges"*)
Pulakos et al. (2000)	No	Measure is individual-focused. Marques-Quinteiro's group-based measure is more appropriate for D11.
Abrantes et al. (2018)	No	Includes ten items. Consistent with D11 working definition. Team improvised adaptation (i.e. whether the team actually deals with changes/obstacles) part of the measure could be completed by the team (as self-report) and could also be completed by observers during the task – i.e. after the intervention? The latter could support identification of behavioural markers. Team pre-emptive adaptation (i.e. whether the team plans how they will adapt to obstacles/changes in the environment) items could be relevant to use if employ a pre-mortem intervention. Although this measure *could* be exploited under this study, it is not guided by the theoretical framework from Pulakos et al. (2000) like other measures are (including the Rousseau & Aubé (2020) measure). Therefore, it was not selected. Example items: Preemptive adaptation -- *"The team prepares in advance how to overcome obstacle that might emerge during task performance";* Improvised adaptation – *"The team deals with unanticipated events on the spot"*
Schippers et al. (2007)	No	Includes 34 items of reflexivity (five of which were postulated to tap into adaptation). Too broad a measure that does not align with definition or conceptual framework, as it concerns reflexivity rather than, specifically, adaptation. Moreover, the five items that *do* tap into adaptation are general and do not align with Pulakos et al (2000) framework. Example item: *"In this team agreed upon actions are usually carried out"*
de Jong & de Ruyter (2004)	No	12 items focused on recovery behaviour in a service industry/business. Not easily exploitable in C2 context as it is set in a business context and would require language to be amended. The measure was developed for this specific study and context and does not provide

Source	Decision	Justification for inclusion/exclusion
		evidence of validity. Also items are tangential to the D11 working definition of TA. Example item: *"In our team we actively seek out areas for continuous improvement of our service recovery"*
Georganta et al. (2020)	No	Includes 12 items. Not a good fit with the D11 working definition as the measure concerns behaviours during task performance but also pre-task plans and post-task reflections/changes made for future tasks. Specifically, it measures (a) plan execution, (b) situation assessment and (c) plan formulation in response to changes. The Rousseau & Aubé (2020) measure is a better fit as it taps specifically into team adaptation (as guided by the Pulakos theoretical framework). Example items: Situation assessment – *"Actively scan the environment to see what is happening might affect your team in the future?"* Plan formulation – *"My team adjusted its task performance strategies in response to changes in the context and progress of the project*" Plan execution – *"Our team worked together in a well-coordinated fashion"*
Rousseau & Aubé (2020)	**Yes**	7-item team-level measure that best taps into the Pulakos framework of all measures reviewed (one item per dimension). Provides one summary mean score of team adaptation (as opposed to multiple mean scores for each individual dimension) and so would be straightforward to exploit and interpret. Example item: *"Members of this team take effective action to deal with crises or urgent work demands"*

Measuring & Training Adaptive Team Performance for improved sensemaking

Previous work (Hutton, et al 2017) has noted that the US DOD has already identified adaptive performance as critical and has provided a mandate (of sorts) through the Army Learning Concept 2015, however, they are still exploring what adaptive performance is, how to train it and how to assess it. Much of this work is not uncomplicated (Ward et al., 2016, 2018). Across the US and UK there appears to be few tangible training interventions or training solutions, performance measures, evidence for efficacy, and support to the training community, to evaluate the efficacy of any new programmes - which continues to provide a challenge for researchers and practitioners in this area and had done so over the past two decades.

DISCUSSION & CONCLUSION

The intended impact of this work is to support guidance decisions for training and education in the area of improving team adaptive performance. We are continuing to explore this area of cognitive and practical complexity.

A number of gaps however, in the empirical evidence to support the effectiveness of existing training approaches aimed specifically at adaptive performance and sensemaking continue to exist. Nevertheless, there is an appetite in many organizations, in health, business, sport and especially in high relability environments to continue to improve our development and understanding of adaptive individuals, teams and sensemaking in the future. Crucially, this study concurs with previous studies on adaptive performance and continues to suggest that there is sufficient scientific and practitioner interest to warrant further development and testing of measures of adapitivity in a range of organisational contexts.

REFERENCES

Abrantes, A. C. M., Passos, A. M., Cunha, M. P., & Miner, A. S. (2020). Managing the unforeseen when time is scarce: how temporal personality and team improvised adaptation can foster team performance. *Group Dynamics: Theory, Research, and Practice, 24*(1), 42-58.

Baard, S. K., Rench, T. A., & Kozlowski, S. W. (2014). Performance adaptation: A theoretical integration and review. *Journal of Management, 40*(1), 48-99.

Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology, 91*(6), 1189-1207.

Clerici, A., Hillyer, C., McEwan, D. and Gore, J. (2020). *Improving team performance adaptation to support sensemaking.* Scoping Study Report D11-D02. Dstl 1000155042

Cooper-Chapman, C. (2018). *Collaborative Decision Making Environments (CDME) Summary Report*. Dstl, Porton Down (GB), DSTL/CR106533 V1.0.

Dalenberg, S., Ad Vogelaar, L. W., & Beersma, B. (2009). The effect of a team strategy discussion on military team performance. *Military Psychology, 21*(sup2), S31-S46.

McEwan, D., Harden, S. M., Zumbo, B. D., Sylvester, B. D., Kaulius, M., Ruissen, G. R., Dowd, A. J., & Beauchamp, M. R. (2016). The effectiveness of multi-component goal setting interventions for changing physical activity behaviour: A systematic review and meta-analysis. *Health Psychology Review*, 10(1), 67-88.

Gore, J. Ward, P, Conway, G., Ormerod, T., Wong, W. & Stanton N. (Eds) (2018). Naturalistic Decision Making:Navigating Uncertainty in Complex Sociotechnical Work *Special Issue Cognition Technology & Work.* 20 (4) 521-527

Gore, J. & Ward, P. (Eds) (2018). Naturalistic Decision Making and Macrocognition Under Uncertainty Theoretical & Methodological Developments. *Journal of Applied Memory & Cognition*. 7 (1) 33-34.

Gore, J., Ward, P., Hutton, R. J. B., Hoffman, R. R., Anderson, T. , Conway, G. (2017). Uncovering adaptive expertise: an interpretive synthesis review. *Reconnecting Management Research within the disciplines: Shaping the Research Agenda for the Social Sciences. British Academy of Management Conference*, University of Warwick.

Gore, J. and Ward, P. (2017) Eds. *Naturalistic Decision Making and Uncertainty. Proceedings of the 13th Bi-Annual Naturalistic Decision Making Conference* University of Bath, UK. ISBN 978-0-86197-194-7

Hutton, R. J. B., Ward, P., & Capewell, D. (2017). DHCSTC TIN 3.234 *Preparation for Decision Making in Complex & Chaotic Situations Final Report Adaptive Skills Training for the Standing Joint Force*

Headquarters at Exercise Joint Venture 2017. BAE Systems, Yeovil (GB), O-DHCSTC-I309549_PT_T3_234/010 V3.

Hyllengren, P. (2017). Military leaders' adaptability in unexpected situations. *Military Psychology, 29*(4), 245-259.

Klein, G., & Pierce, L. (2001). Adaptive teams. *Proceedings of 6th International Command and Control Research and Technology Symposium, 46*, 225-229.

Marques-Quinteiro, P., Ramos-Villagrasa, P. J., Passos, A. M., & Curral, L. (2015). Measuring adaptive performance in individuals and teams. *Team Performance Management, 21*(7/8), 339-360.

Marques-Quinteiro, P., Rico, R., Passos, A. M., & Curral, L. (2019). There is light and there is darkness: On the temporal dynamics of cohesion, coordination, and performance in business teams. *Frontiers in Psychology, 10*, 24 April 2019.

Maynard, M. T., Kennedy, D. M., & Sommer, S. A. (2015). Team adaptation: A fifteen-year synthesis (1998–2013) and framework for how this literature needs to "adapt" going forward. *European Journal of Work and Organizational Psychology, 24*(5), 652-677.

Oprins, E. A. P. B., van den Bosch, K., & Venrooij, W. (2018). Measuring adaptability demands of jobs and the adaptability of military and civilians. *Military Psychology, 30*(6), 576-589.

Ployhart, R. E., Saltz, J. L., Mayer, D. M., & Bliese, P. D. (2002). Individual Adaptability: Measurement, Construct Validity, and Relations to Leadership Performance. Paper presented at the *Annual Conference of the International Personnel Management Association Assessment Council*, New Orleans, LA.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612.

Rico, R., Gibson, C. B., & Sánchez-Manzanares, M. (2020). Team adaptation and the changing nature of work: lessons from practice, evidence from research, and challenges for the road ahead. *Australian Journal of Management, 45*(3), 507-526.

Roberts, A., Webster, L., Pope, K., Fay, D., Twist., & Fenton, C. (2019). *DHCSTC TIN 2.108: A Review of Teamwork Measures and Interventions.* O-DHCSTC_I2_T_T2_108_003. V3.

Rosen, M. A., Bedwell, W. L., Wildman, J. L., Fritzsche, B. A., Salas, E., & Burke, S. (2011). Managing adaptive performance in teams: Guiding principles and behavioral markers for measurement. *Human Resource Management Review, 21*, 107-122.

Rousseau, V., & Aubé, C. (2020). Disentangling the relationship between empowering leader behaviors and adaptive performance in work teams. *Group Processes & Intergroup Relations, 23*(5), 761-777.

Salas E., Nichols, D. R., & Driskell, J. E. (2007). Testing three team training strategies in intact teams: a meta-analysis. *Small Group Research, 38*(4), 471-488.

Stanton, N., & Roberts, A. (2018). *DHCSTC TIN 3.113 Command Team Experimental Testbed (ComTET): Understanding Submarine Command Team Functionality from a Sociotechnical Perspective*. BAE Systems, Yeovil (GB). O-DHCSTC_I389205_H_T3_113/004 V2.

Tucker, J. S., & Gunther, K. M. (2009). The application of a model of adaptive performance to army leader behaviors. *Military Psychology, 21*(3), 315-333.

Ward, P., Hutton, R., Hoffman, R., Gore, J., Anderson, T., & Leggatt, A. (2016). *DHCSTC TIN 2.077 - Developing Skilled Adaptive Performance: A Scoping Study*. BAE Systems, Yeovil (GB), O-DHCSTC_I2_T_T2_077/005 V3.

Ward, P., Gore, J. Hutton, R. J. B., Conway, G. & Hoffman, R. R. (2018). Adaptive Skill as The Sine Qua Non of Expertise. *Journal of Applied Memory and Cognition*. 7 (1) 35-50.

Ward, P., Schraagen, JM, Gore, J. & Roth, E. (2020). *The Oxford Handbook of Expertise*. Oxford University Press

White, S. S., Mueller-Hanson, R. A., Dorsey, D. W., & Pulakos, E. D., Wisecarver, M. M., Deagle, E. A., & Mendini, K. G. (2005). *Developing Adaptive Proficiency in Special Forces Officers (Research Report 1831).* U.S. Army Research Institute for the Behavioral and Social Sciences.

Rethinking mass casualty distribution – Embedding a resilient hospital selection algorithm into a mass casualty distribution simulation model

Sheuwen Chuang[a], Chia-Hsin Cheng[a], Ching-An Lee[b]
[a]*Graduate Institute of Data Science, Taipei Medical University, Taipei, Taiwan*
[b]*Fire Department, New Taipei City, Taiwan*

ABSTRACT

Considering hospitals' resilience potential in addition to the shortest-distance and medical-care adequacy policy to distribute mass casualties (MC) is an imperative practice for mass casualty distribution decision making. This study developed a novel hospital selection algorithm composed of driving time from the disaster site to hospital, care adequacy, and mobilization ability to determine the best hospital choice for MC distribution. Next, we developed an MC distribution simulation model embedding the algorithm to generate optimized distribution decisions for various MC incident scenarios. The simulation model was tested by using the Formosa Fun Coast Dust Explosion. Regarding super overload on some responsible emergency hospitals in the FFCDE event, the model with mobilization ability shows a better-balanced distribution of mass casualties to the initial receiving hospitals than without the ability. The study findings can contribute to surge capacity planning and resource assessment of emergency medical services for future disasters.

KEYWORDS

Decision making; Emergency Medical Service; Resilience, Mass Casualty Distribution; Discrete-event simulation.

INTRODUCTION

Large-scale events can produce casualties that threaten medical care systems. Decision making on the distribution of mass casualties to appropriate hospitals is a critical practice in mass casualty incident (MCI) management. Depending on the magnitude of the event, hospitals with different levels of surge capacity following a mass casualty incident fall into varying degrees of overwhelmed by the influx of patients, impacting the hospital's functional operations, and consequently on the safety of casualties.

In general, there are two principles of hospital selection for MC distribution, which include: the shortest distance from the disaster scene to individual hospitals and medical care adequacy accordingly for the received patients (Emergency Medical Services Act, 2013). The first principle is convenient to apply by using the information from the internet. The second principle relies on static government data about hospital medical care capacity. However, these data do not tell the hospitals' resilience potential, such as the mobilization ability to extend their emergency care capability. Disasters seem to become increasingly common in current decades. To better prepare emergency care service staff to offer patients timely and appropriate care for survival in various challenging situations during or after disasters, it is critical to learn how to distribute mass casualties and deploy resources for emergency medical care.

The purpose of the study is first to develop a hospital selection algorithm composed of driving time from the disaster scene to individual hospitals, care adequacy, and mobilization ability to determine the best hospital choice for MC distribution. Next, embedding the algorithm to develop a mass casualty distribution simulation model to generate optimized MC distribution decisions for various MC incident scenarios.

METHODS

The algorithm development collected data, including hospital capacity in beds, staff, and accreditation results for the emergency responsibility hospital (ERH) from the open government data (MOHW, 2017; MOHW, 2020) as the input for computing the hospital's care adequacy and mobilization ability, followed by the calibration of the algorithm parameters based on the interview data with 36 key participants from the emergency operation centers and five initial recurving hospitals (Chuang, Woods, Ting, et al., 2019; Chuang, Woods, Lee, et al., 2018; Chuang, Chang, Woods, et al., 2018). Hospital's address is used to extract the shortest driving time from a disaster scene via the Google Map API. Total 205 emergency responsibility hospitals in Taiwan were collected.

Modeling the mass casualty distribution system used the discrete event simulation technique in the Anylogic software. The simulation model was developed into eight modules (Figure 1): the initiation module of a mass casualty incident, the data module of setting parameters in the user interface and import data, and the module of rescue, triage, waiting for ambulances, hospital selection, distribution, and export data.

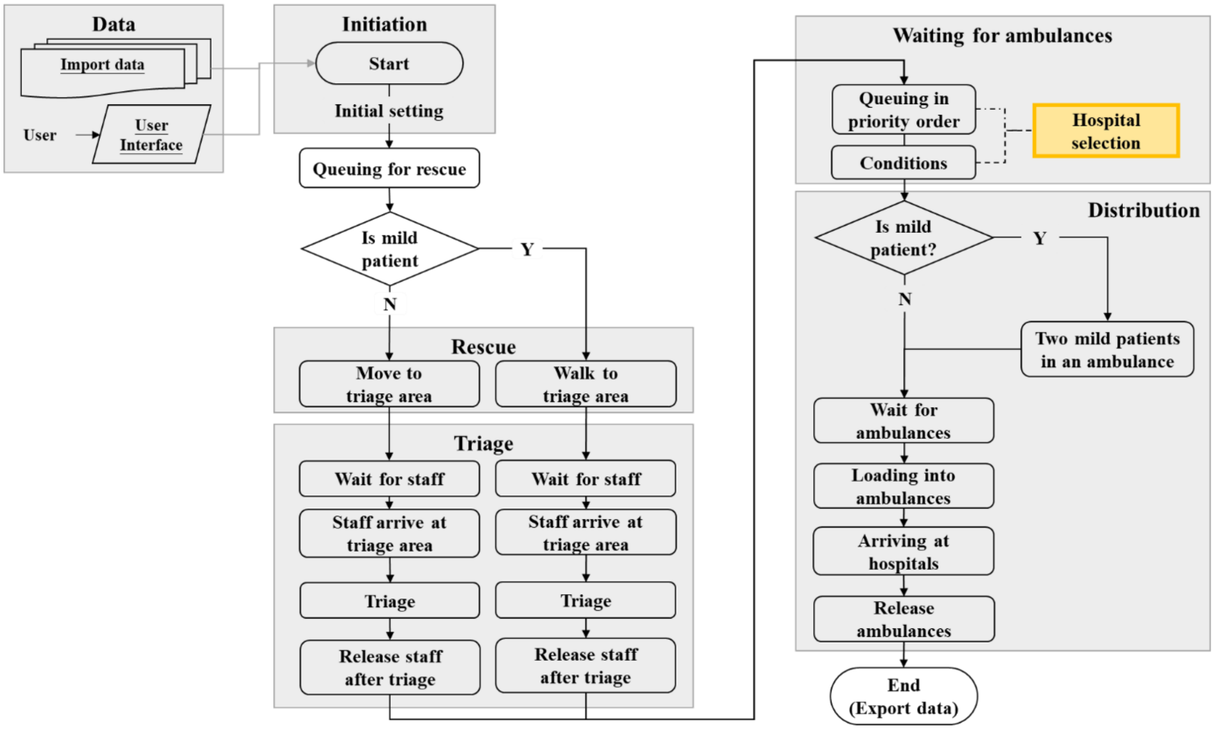


Figure 1. The simulation model framework

The Formosa Coast Dust Explosion (FFCDE) event was adopted to test the model. The FFCDE event occurred on Jan 27, 2015 at Color Play Asia water part party in New Taipei City, Taiwan. The flammable properties of the swimwear in which attendants were dressed resulted in large total body surface area burns (TBSA, average 44 %; 281 people with TBSA > 40%, 41 people > 80%). One bus and 144 regular ambulances were deployed to the field. A total of 301 (60.3%) patients were distributed to hospitals via ambulance; others were self-transported to hospitals. Within 6 hours, 499 burn victims had been transported to 36 hospitals including 10 large medical centers, 23 regional hospitals, and 3 district hospitals across regions (MOHW, 2017).

RESULTS

Hospital Selection algorithm

The hospital selection algorithm is to prioritize hospitals identified in the area to support mass casualty distribution decision making. A hospital selection algorithm was developed, which hospital score $Score_{ij}$ is an equation constructed by three elements in equation (1): normalized driving time, normalized care adequacy adjusting by patient load on hospitals, and normalized mobilization ability.

$$Score_{ij} = x_{ij} + y_{ij}(1 - z_{ij}) + y_{ij}w2_{ij}$$

$$= x_{ij} + y_{ij}\left[1 - \frac{\sum_{i=1}^{3} w1_{ij} \times RC_{ij}}{ED\ Beds_j}\right] + y_{ij} \times \frac{Total\ Medical\ Staff_j}{Critical\ Beds_j} \quad (1)$$

Where i =1, 2, 3 indicates the patient severity, the i =1 indicates mild, the i =2 indicates moderate, the i =3 indicates severe.

Where j=1, 2, 3, …, n is the individual hospital number identified in the area to support mass casualty incident.

The driving time x_i is from the disaster scene to each hospital, and the y_{ij} is a digital number representing care adequacy shown in Table 1. Based on the hospital accreditation results, each hospital is accredited as a specific medical care capacity. However, only the three primary levels of ERH (1: severe level, 2: medium level, 3: general level in Table 1) are utilized by the local emergency operation centers for mass casualty distribution. The study adopts the additional six sub-levels to create a new two-digit care capacity code corresponding to the patient's triage level. Besides, hospital care adequacy was dynamic changing as the patient surge arrived. Thus, the y_{ij} has to be adjusted accordingly by the penalty parameter z_{ij}. The z_{ij} is the quotient of summation of weighting factor $w1_{ij}$ multiplying received casualty RC_{ij} divided by the $ED\ Bed_j$. The penalty z_{ij} increases as the hospital receives more patients, and the higher the penalty parameter, the lower the care adequacy. Based on the interview data, the more staff hospital has, the more supporting workforce they can mobilize.

The mobilization weight $w2_{ij}$ is determined by the ratio of total medical staff *Total Medical Staff$_j$* in a hospital to the amount of critical beds *Critical Beds$_j$*. It means the number of staff per critical bed that a hospital can utilize, the higher the ratio is, the higher the mobilization ability. The total medical staff sum of number of staff in the surgery, emergency department, anesthesia, plastic surgery, neurosurgery, oral pathology, oral maxillofacial, internal medicine, general western medicine, pediatrics, radiation therapy, neurology, orthopedics. The critical beds sum of acute beds, ICU beds, *ED Bed$_j$*, hospitalist beds. The mobilization weight reflects the mobilization ability of a hospital.

The hospital score $Score_{ij}$ is the sum of three major parameters. In the simulation model, each patient has a list of scores for all possible sending hospitals according to their triage level. The hospital would have a higher score if the x_{ij} is less, y_{ij} is higher, z_{ij} is lower, $w2_{ij}$ is higher. The highest hospital score responding to a patient indicates the best choice hospital for the patient in the perspective of global optimization of mass casualty distribution.

Table 1. Medical-care adequacy matrix

Emergency Responsibility Hospital (ERH) level	EOC adopted Primary ERH levels	y_severe	y_moderate	y_mild
General	1	10	10	30
General + Medium EM	1	11	11	31
General + Medium ICU	1	11	11	31
General + Medium EM, PN	1	12	12	32
General + Medium EM, MT	1	12	12	32
General + Medium EM, ICU	1	12	12	32
General + Medium EM, AS, ICU	1	13	13	33
General + Medium EM, STEMI, ICU	1	13	13	33
General + Medium EM, AS, MT	1	13	13	33
General + Medium EM, AS, STEMI, ICU	1	14	14	34
General + Medium EM, AS, PN, ICU	1	14	14	34
General + Medium EM, MT, ICU	1	13	13	33
Medium Temp	2	20	30	20
Medium	2	20	30	20
Medium*	2	20	30	20
Medium* + Severe AS	2	21	31	21
Medium* + Severe STEMI	2	21	31	21
Medium* + Severe AS, STEMI	2	22	32	22
Medium* + Severe EM, AS, STEMI	2	23	33	23
Medium + Severe AS	2	21	31	21
Medium + Severe STEMI	2	21	31	21
Medium + Severe PN	2	21	31	21
Medium + Severe AS, STEMI	2	22	32	22
Medium + Severe AS, STEMI, ICU	2	23	33	23
Medium + Severe EM, AS, STEMI, PN, ICU	2	25	35	25
Severe	3	30	20	10

Note: AS: Acute Stroke, STEMI: ST-Elevation Myocardial Infarction, PN: High risk Pregnancy and Neonates, MT: Major Trauma, EM: Emergency Medicine, ICU: Intensive Care Unit.
Any of y_severe, y_moderate, y_mild is a two-digit code representing hospital care adequacy corresponding to a patient's triage level in the perspective of optimal care. Source: MOHW, 2020.

Total transportation time and number of initial receiving hospitals

The model shows less total transportation time for distributing 301 patients than in the event (Table 2). The number of initial receiving hospitals (IRHs) in simulation has 18 more hospitals than in the actual event.

Table 2. Comparison of transportation time and resources utilized between the FFCDE event and the simulation model

Number of patients	Categories			FFCDE	Model	Difference
301 (247 severe patients, 159 moderate patients, 27 mild patients)	Transportation time for all patients by ambulance depart from the disaster scene			6 hrs.	4 hr. 41 mins	-1 hr. 19 mins
	Ambulance	Number of ambulances		144*	144	0
		Number of trips	1 trip	*	2	-
			2 trips	*	140	-
			3 trips	*	2	-
495**	ERH level	Severe Hospital		19	17	-2
		Medium Hospital		13	20	+7
		General Hospital		4	17	+13
		Total		36	54	+18

** In the FFCDE event, only number of ambulances were recorded, and buses were used to transport patients.*

*** 495 patients are used only to compare number of initial receiving hospitals between the FFCDE event and the model due to about 40% of mass casualties were self-transported.to hospitals.*

Table 3 shows more initial receiving hospitals were identified in the simulation model than in the FFCDE event, and severe-level ERHs received more moderate and mild patients in the FFCDE than in the simulation, and more moderate and mild patients were sent to medium-level ERHs in the simulation model than the FFCDE.

Table 3. Number of patients received by triage level corresponding to ERH level

ERH level	Severe patients FFCDE/model (gap)		Moderate patients FFCDE/model (gap)		Mild patients FFCDE/model (gap)		Total IRHs FFCDE/model (gap)	
Number of initial receiving hospitals (IRHs)								
Severe Hospital	17 / 15	(-2)	16 / 13	(-3)	12 / 1	(-11)	19 / 17	(-2)
Medium Hospital	10 / 4	(-6)	11 / 19	(+8)	7 / 7	(0)	13 / 20	(+7)
General Hospital	1 / 2	(+1)	3 / 4	(+1)	2 / 14	(+12)	4 / 17	(+13)
Total	28 / 21	(-7)	30 / 36	(+6)	21 / 22	(+1)	36 / 54	(+18)
Number of patients								
Severe Hospital	195 / 208	(+13)	111 / 46	(-65)	67 / 4	(-63)	373 / 258	(-115)
Medium Hospital	45 / 28	(-17)	42 / 97	(+55)	16 / 26	(+10)	103 / 151	(+48)
General Hospital	7 / 11	(+4)	6 / 16	(+10)	6 / 59	(+53)	19 / 86	(+67)
Total	247		159		89		495	

Note: ERH: Emergency Responsibility Hospital.

Comparison of number of patients received by IRHs in different characteristics

Table 4 shows that severe patients were most likely sent to public medium-level hospitals in the FFCDE event than in the model, and moderate patients were most likely sent to private medium-level hospitals in the model. Mild patients were likely sent to private severe-level hospitals in the FFCDE event.

Table 4. Comparison of number of patients received between public and private hospitals

ERH level, ownership	Severe patients FFCDE/model (gap)		Moderate patients FFCDE/model (gap)		Mild patients FFCDE/model (gap)		Total patients FFCDE/model (gap)	
Severe Hospital								
Public Hospital	61 / 60	(-1)	39 / 13	(-26)	19 / 0	(-19)	119 / 73	(-46)
Private Hospital	134 / 148	(+14)	72 / 33	(-39)	48 / 4	(-44)	254 / 185	(-69)
Medium Hospital								
Public Hospital	36 / 22	(-14)	31 / 40	(+9)	5 / 14	(+9)	72 / 76	(+4)
Private Hospital	9 / 6	(-3)	11 / 57	(+46)	11 / 12	(+1)	31 / 75	(+44)
General Hospital								
Public Hospital	7 / 2	(-5)	2 / 7	(+5)	0 / 12	(+12)	9 / 21	(+12)
Private Hospital	0 / 9	(+9)	4 / 9	(+5)	6 / 47	(+41)	10 / 65	(+55)
Total patients	247		159		89		495	

Note: ERH: Emergency Responsibility Hospital.

Although the total number of initial receiving serve-level ERHs is a minor difference between the FFCDE event and the model, there is a large variance of 115 patients sent to the severe-level ERHs between these two (Table 5). The highest difference in total patients between the FFCDE event and the model occurs in hospital S01. S01 is the nearest severe-level ERH to the disaster scene. The real case sent 26 more patients to the hospital than in the simulation model, which included 21 moderate patients. As a result, the emergency department of hospital S01 was super overload and had to break through the challenges beyond its surge capacity (Chuang, Chang, Woods, et al., 2018). The second-highest difference is in hospital S13 (a public hospital). The hospital received 15 more moderate patients and 10 more mild patients in the FFCDE event than in the model. The third-highest difference is in hospital S04. In general, mass casualties were distributed more balanced in the model than in the real event.

Table 5. Comparison of number of patients received between severe-level ERHs

Severe -level ERHs	Driving Time (mins)	Mobilization ability (w2)	FFCDE / Model (gap) 19 / 17 (-2)			
			Severe patients FFCDE/model (gap)	Moderate patients FFCDE/model (gap)	Mild patients FFCDE/model (gap)	Total patients FFCDE/model (gap)
S01	25	0.7045	17 / 20 (+3)	24 / 3 (-21)	8 / 0 (-8)	49 / 23 (-26)
S02	27	0.4183	24 / 36 (+12)	13 / 5 (-8)	8 / 0 (-8)	45 / 41 (-4)
S03	30	0.7172	14 / 26 (+12)	7 / 3 (-4)	7 / 0 (-7)	28 / 29 (+1)
S04	30	0.5616	27 / 9 (-18)	1 / 0 (-1)	3 / 0 (-3)	31 / 9 (-22)
S05	30	0.4322	9 / 18 (+9)	2 / 3 (+1)	2 / 0 (-2)	13 / 21 (+8)
S06	31	0.8841	44 / 30 (-14)	2 / 5 (+3)	2 / 0 (-2)	48 / 35 (-13)
S07	31	0.5248	8 / 8 (0)	2 / 1 (-1)	10 / 0 (-10)	20 / 9 (-11)
S08	32	0.5866	5 / 14 (+9)	10 / 3 (-7)	5 / 0 (-5)	20 / 17 (-3)
S09	32	0.4226	13 / 16 (+3)	0 / 4 (+4)	-	13 / 20 (+7)
S10	34	0.2408	2 / 5 (+3)	5 / 0 (-5)	2 / 0 (-2)	9 / 5 (-4)
S11	35	0.4562	2 / 8 (+6)	12 / 6 (-6)	-	14 / 14 (0)
S12	35	0.2733	1 / 4 (+3)	1 / 1 (0)	-	2 / 5 (+3)
S13	36	0.4635	10 / 11 (+1)	20 / 5 (-15)	10 / 0 (-10)	40 / 16 (-24)
S14	38	0.2298	6 / 1 (-5)	3 / 0 (-3)	6 / 0 (-6)	15 / 1 (-14)
S15	39	0.5634	4 / 2 (-2)	0 / 3 (+3)	-	4 / 5 (+1)
S16	40	0.5620	8 / 0 (-8)	6 / 4 (-2)	-	14 / 4 (-10)
S17	42	0.2273	-	2 / 0 (-2)	-	2 / 0 (-2)
S18	43	0.5279	1 / 0 (-1)	-	0 / 4 (+4)	1 / 4 (+3)
S19	60	0.2534	-	1 / 0 (-1)	4 / 0 (-4)	5 / 0 (-5)
Total patients			195 / 208 (+13)	111 / 46 (-65)	67 / 4 (-63)	373 / 258 (-115)

Note: ERHs: Emergency Responsibility Hospitals.

There is a difference of 55 moderate patients sent to the medium-level ERHs between the FFCDE event and the simulation model (Table 6). The highest difference (+18) occurs in hospital M03, zero patient was sent to the hospital, but 6 severe patients and 12 moderate patients were sent to the hospital in the model. M03 is relatively a new medium-level private ERH comparing to other ERHs in the same area. MO2, M05 are the public ERHs in the area, and they received either more severe or moderate patients in the FFCDE event than in the model. Overall, five additional medium-level hospitals are included to receive patients in the model.

Table 6. Comparison of number of patients received between medium-level ERHs

Medium -level ERHs	Driving Time (mins)	Mobilization ability (w2)	FFCDE / Model (gap) 13 / 20 (+7)			
			Severe patients FFCDE/model (gap)	Moderate patients FFCDE/model (gap)	Mild patients FFCDE/model (gap)	Total patients FFCDE/model (gap)
M01	22	0.2393	3 / 12 (+9)	5 / 2 (-3)	3 / 0 (-3)	11 / 14 (+3)
M02	25	0.1989	7 / 8 (+1)	20 / 5 (-15)	1 / 0 (-1)	28 / 13 (-15)
M03	26	0.2554	0 / 6 (+6)	0 / 12 (+12)	-	0 / 18 (+18)
M04	27	0.2015	0 / 2 (+2)	0 / 10 (+10)	-	0 / 12 (+12)
M05	29	0.1862	15 / 0 (-15)	0 / 6 (+6)	1 / 0 (-1)	16 / 6 (-10)
M06	32	0.2529	-	0 / 5 (+5)	-	0 / 5 (+5)
M07	32	0.2442	4 / 0 (-4)	2 / 7 (+5)	-	6 / 7 (+1)
M08	35	0.3846	-	0 / 10 (+10)	-	0 / 10 (+10)
M09	35	0.3014	-	0 / 4 (+4)	-	0 / 4 (+4)
M10	35	0.1549	4 / 0 (-4)	2 / 4 (+2)	-	6 / 4 (-2)
M11	36	0.1453	1 / 0 (-1)	1 / 4 (+3)	-	2 / 4 (+2)
M12	36	0.2500	2 / 0 (-2)	1 / 11 (+10)	3 / 0 (-3)	6 / 11 (+5)
M13	36	0.2050	3 / 0 (-3)	1 / 4 (+3)	-	4 / 4 (0)
M14	38	0.1076	4 / 0 (-4)	7 / 6 (-1)	4 / 2 (-2)	15 / 8 (-7)
M15	41	0.1522	-	1 / 2 (+1)	0 / 2 (+2)	1 / 4 (+3)
M16	42	0.0890	-	0 / 1 (+1)	0 / 4 (+4)	0 / 5 (+5)
M17	42	0.2527	-	0 / 2 (+2)	2 / 4 (+2)	2 / 6 (+4)
M18	46	0.2260	2 / 0 (-2)	1 / 1 (0)	2 / 4 (+2)	5 / 5 (0)
M19	46	0.1667	-	0 / 1 (+1)	0 / 4 (+4)	0 / 5 (+5)
M20	51	0.2330	-	1 / 0 (-1)	0 / 6 (+6)	1 / 6 (+5)
Total patients			45 / 28 (-17)	42 / 97 (+55)	16 / 26 (+10)	103 / 151 (+48)

Note: ERHs: Emergency Responsibility Hospitals.

In Table 7, the model found extra 16 moderate patients sent to the medium-level ERHs. The general-level ERHs received 53 more mild patients in the model than in the FFCDE event; additional 13 general-level ERHs that their driving time to the disaster scene are within 25 – 66 minutes were included in the model to receive mild patients.

Overall, regarding the hospitals located within 35 minutes driving time to the disaster scene (Table 5-7), there are 30 hospitals, including 12 severe-level ERHs, ten medium-level ERHs, and eight general-level ERHs. Fifty-eight mild patients were sent these hospitals, including 47 sent to 9 severe-level, five sent to 2 medium-level, and six sent to 2 general-level in the FFCDE event. However, only 31 mild patients were distributed to 5 general-level ERHs in this area in the model.

Table 7. Comparison of number of patients received between general-level ERHs

General-level ERHs	Driving Time (mins)	Mobilization ability (w2)	FFCDE / Model (gap) 13 / 20 (+7)							
			Severe patients FFCDE/model (gap)		Moderate patients FFCDE/model (gap)		Mild patients FFCDE/model (gap)		Total patients FFCDE/model (gap)	
G01	25	0.3333	0 / 9	(+9)	0 / 2	(+2)	-		0 / 11	(+11)
G02	27	0.2393	7 / 2	(-5)	2 / 7	(+5)	-		9 / 9	(0)
G03	29	0.2411	-		3 / 4	(+1)	5 / 0	(-5)	8 / 4	(-4)
G04	33	0.2828	-		0 / 3	(+3)	0 / 16	(+16)	0 / 19	(+19)
G05	33	0.0870	-		-		1 / 4	(+3)	1 / 4	(+3)
G06	34	0.2462	-		-		0 / 2	(+2)	0 / 5	(+2)
G07	34	0.2045	-		-		0 / 5	(+5)	0 / 5	(+5)
G08	34	0.1048	-		-		0 / 4	(+4)	0 / 4	(+4)
G09	36	0.1096	-		-		0 / 4	(+4)	0 / 4	(+4)
G10	36	0.1087	-		-		0 / 2	(+2)	0 / 2	(+2)
G11	40	0.1333	-		-		0 / 4	(+4)	0 / 4	(+4)
G12	41	0.2143	-		-		0 / 2	(+2)	0 / 2	(+2)
G13	41	0.1420	-		-		0 / 4	(+4)	0 / 4	(+4)
G14	43	0.1532	-		-		0 / 4	(+4)	0 / 4	(+4)
G15	44	0.2197	-		-		0 / 4	(+4)	0 / 4	(+4)
G16	47	0.2174	-		-		0 / 2	(+2)	0 / 2	(+2)
G17	66	0.2830	-		-		0 / 2	(+2)	0 / 2	(+2)
G18	35	*	-		1 / 0	(-1)	-		1 / 0	(-1)
Total patients			7 / 11	(+4)	6 / 16	(+10)	6 / 59	(+53)	19 / 86	(+67)

Note: ERHs: Emergency Responsibility Hospitals.
**The hospital G18 withdrew from the ERH on Mar. 1, 2017, but it was ERH at the FFCDE event in 2015 so that can not collect the hospital information data.*

Comparison between the algorithm with and without mobilization ability

While considering mobilization ability (w2), the severe-level ERHs can receive more severe patients in the model with w2 than without adopting mobilization ability. The medium-level ERHs can receive less severe patients but more moderate patients in the w2 model than without the w2 model (Table 8). However, the mobilization ability in the algorithm for hospital selection did not apparently affect the number of initial receiving hospitals.

Table 8. Difference between the algorithm with and without mobilization ability

ERH level	Severe patients without w2 model/ with w2 model (gap)		Moderate patients without w2 model/ with w2 model (gap)		Mild patients without w2 model/ with w2 model (gap)		Total IRHs without w2 model/ with w2 model (gap)	
<u>Number of initial receiving hospitals (IRHs)</u>								
Severe Hospital	14 / 15	(+1)	12 / 13	(+1)	4 / 1	(-3)	18 / 17	(-1)
Medium Hospital	5 / 4	(-1)	19 / 19	(0)	9 / 7	(-2)	20 / 20	(0)
General Hospital	3 / 2	(-1)	8 / 4	(-4)	14 / 14	(0)	17 / 17	(0)
Total	22 / 21	(-1)	39 / 36	(-3)	27 / 22	(-5)	55 / 54	(-1)
<u>Number of patients</u>								
Severe Hospital	162 / 208	(+46)	54 / 46	(-8)	12 / 4	(-8)	228 / 258	(+30)
Medium Hospital	61 / 28	(-33)	79 / 97	(+18)	30 / 26	(-4)	170 / 151	(-19)
General Hospital	24 / 11	(-13)	26 / 16	(-10)	47 / 59	(+12)	97 / 86	(-11)
Total	247		159		89		495	

Note: ERH: Emergency Responsibility Hospital.

DISCUSION AND CONCLUSION

The study's findings highlight mass casualty distribution in an actual event could produce super overload on some initial receiving hospitals with the characteristics in nearby the disaster scene, public, and severe-level or medium-level ERHs. The sudden on-set event created extreme demand beyond their surge capacity regularly prepared. The initial receiving hospitals were pushed into a position that required them to develop additional surge capacity to provide care to the mass casualties. Fortunately, they succeed in providing emergency medical care to the patients in the FFCDE event. However, suppose the emergency medical services units and staff can exercise using the simulation model. In that case, they can learn how to effectively and efficiently respond to different scenarios in mass casualty distribution and preparedness planning.

Considering hospitals' mobilization potential for additional internal and external capacity, in addition to the registered capacity, which can extend hospitals' care capability toward more resilience. The simulation model adopts the resilience concepts and the successful hospitals' experiences performed in the FFCDE event to develop a novel hospital selection algorithm for supporting mass casualty distribution decision making. The algorithm is constructed by three elements: normalized driving time, normalized care adequacy adjusting by patient loading on hospitals, and normalized mobilization ability. The study shows the simulation model embedding the algorithm can produce an optimal mass casualty distribution results, including balanced distribution of mass casualties to the initial receiving hospitals, and less total transportation time. However, the algorithm may still need to be further advanced for handling dynamic changes of load and capacity. Using the simulation model can contribute to the resource assessment of emergency medical services and hospital surge capacity planning in different disaster scenarios.

ACKNOWLEDGMENTS

We thank all authors for their contributions to the paper.

REFERENCES

Chuang, S., Woods, D.D., Lee, C.-A., Cheng, C.-H., Chen, H.-C. (2018). Coping with communication challenges after the Formosa Fun Coast Dust Explosion, *Proceedings of the 2018 Resilience Week (RWS)* (pp. 5-10), Denver, USA. The Institute of Electrical and Electronics Engineers (IEEE). doi: 10.1109/rweek.2018.8473499

Chuang, S., Chang, K.-S., Woods, D.D., Chen, H.-C., Reynolds M.E., Chien D.K. (2018). *Beyond surge: Coping with mass burn casualty in the closest hospital to the Formosa Fun Coast Dust Explosion*. Burns, 45(4):964-973. doi: 10.1016/j.burns.2018.12.003

Chuang, S., Woods, D.D., Ting H.W., Cook, R.T., Hsu, J.-C., (2019). *Coping with Mass Casualty: Insights into a Hospital's Emergency Response and Adaptations After the Formosa Fun Coast Dust Explosion*. Disaster Med Public Health Prep, 14(4), 467-476. doi: 10.1017/dmp.2019.69

Emergency Medical Services Act 1995, Ministry of Health and Welfare, Taiwan. §5. (1995 & rev. 2013). Retrieved Mar 31, 2020 from https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0020045

Emergency Operation Center, New Taipei City. (2015). *519 injured in the Formosa Coast Dust Explosion event, 184 people in the intensive care unit (list included)*. Retrieved Jan 20, 2019 from https://news.ltn.com.tw/news/society/breakingnews/1361682 [In Chinese]

Lee, C.-A. (2015). *627 Formosa Coast Dust Explosion disaster response and recovery experience sharing*. Retrieved Jun 24, 2019 from http://ws.ndc.gov.tw/001/administrator/10/ckfile/4e9135d7-edd7-4de5-885d-e76bdd8abf65.pdf [In Chinese]

Ministry of Health and Welfare. (2017). *Save lives at all costs: disaster response to the 2015 Formosa Fun Coast Dust Explosion by the Ministry of Health and Welfare*. Retrieved Jan 10, 2021 from https://www.mohw.gov.tw/dl-14512-cba8d50f-6a68-47ad-baec-5799a389c7a2.html [In Chinese]

Ministry of Health and Welfare. (2017). *Hospital information platform*. Retrieved Jan 04, 2021 from https://mcia.mohw.gov.tw/openinfo/A100/A101-1.aspx [In Chinese]

Ministry of Health and Welfare. (2020). *1091101 Emergency responsibility hospital accreditation standards*. Retrieved Nov 20, 2020 from https://www.mohw.gov.tw/dl-42934-a39dba78-cda7-48da-bdc7-2659af7d57dd.html [In Chinese]

Taipei City Government. (2020). *Formosa Coast Dust Explosion platform*. Retrieved Jan 20, 2019 from https://beta.gov.taipei/mp2015ffc/Default.aspx [In Chinese]

Wang, T. H., Jhao, W. S., Yeh, Y. H., & Pu, C. (2017). *Experience of distributing 499 burn casualties of the June 28, 2015 Formosa Color Dust Explosion in Taiwan*. Burns, 43(3), 624-631. doi:10.1016/j.burns.2016.10.008.

Yang, C. C., & Shih, C. L. (2016). *A Coordinated Emergency Response: A Color Dust Explosion at a 2015 Concert in Taiwan*. Am J Public Health, 106(9), 1582-1585. doi:10.2105/AJPH.2016.303261.

The Power of Curiosity

Gary Klein
ShadowBox LLC

ABSTRACT
Curiosity is an intellectual force — a motive and an affect — that drives many types of human activities such as discovery, education, and social sophistication. It is important within the Naturalistic Decision Making community because it is central to many macrocognitive functions, particularly sensemaking, problem detection, and anomaly detection. This theoretical paper summarizes the state of understanding of what curiosity is, how it works, what can block it, and tactics for harnessing it for training and the development of expertise.

KEYWORDS
Curiosity, discovery, inquiry, training.

INTRODUCTION

Curiosity is an intellectual force — a motive and an affect — that drives many types of human activities such as discovery, education, and social sophistication. It is important within the Naturalistic Decision Making community because it is central to many macrocognitive functions, particularly sensemaking, problem detection, and anomaly detection.

My own interest in curiosity was initiated by discussions with highly experienced training supervisors. One of them told me that he had hired some retired personnel to conduct courses but all they wanted to do was present material on powerpoints. He wanted them to explore questions with the trainees but they had trouble doing so. His attempts to get them to pose good questions had failed. Then he realized that the problem wasn't that the trainers had trouble posing questions, it was that they had little or no curiosity about what the trainees were thinking, and he didn't know how to encourage curiosity. Another experienced training supervisor complained to me that he knew of a great many ways that instructors squelched curiosity but he didn't know any ways to spark curiosity. These comments provoked me to wonder about the nature of curiosity.

At the same time, my colleague Joseph Borders and I were designing a program to help instructors become more effective and part of that program was to alter their mindsets, from a mindset that criticizes trainees for making mistakes to a mindset that becomes curious about those mistakes. Therefore, if curiosity was going to play an important part of instructor development, we needed to understand it better.

This is a theoretical paper. It summarizes the state of understanding of what curiosity is, how it works, what can block it, and tactics for harnessing it for training and the development of expertise. The primary audience for this paper is the community of teachers and trainers and instructors who can benefit from becoming more effective at harnessing curiosity.

Before examining the definition of curiosity, it may be useful to calibrate our understanding of the topic by listing some of the claims that are often expressed:

1. Once curiosity is aroused it persists until it gets satisfied
2. The larger the gap between what you know and what you want to know, the greater your curiosity
3. The less you know about a topic, the more curious you are
4. Instructors should encourage trainees to ask questions
5. You can inhibit curiosity but you can't generate it

My position is that all of these claims are problematic. Some are wrong, others are overstated. By the end of this paper, I hope to convince readers to view these claims with some degree of skepticism.

DEFINITION OF CURIOSITY

Curiosity is the motivation to bridge a gap. But what kind of gap?

Loewenstein (1994) sees it as an information gap — You perceive a gap in your knowledge or understanding. The gap is between what you know vs what you want to know. (For something to be regarded as novel, complex or ambiguous depends on your existing knowledge and your knowledge goals.)

Within this framework, a primary function of curiosity is to energize us to seek out information that is capable of resolving the gap.

Loewenstein also suggests that curiosity is aroused by sensemaking — which we can regard as a sensemaking gap — we try to find order, to see patterns and intentions. We are curious about events we didn't expect and cannot explain.

Kashdan (2004) sees the gap differently, as a mental model gap. Curiosity arises as we try to integrate new information into our mental model. The more difficulty we have, the greater our curiosity.

My view differs from that of Loewenstein and Kashdan (Klein, 2020). I suggest that there is a plausibility gap. Our efforts to explain events often take the form of a story, a set of state transitions. If any of those transitions are dubious and create doubt and strain, then curiosity arises to clarify the transition. This formulation is about finding a satisfying/plausible frame. The stopping rule of this story-building process is that we need to see the transition as plausible. This plausibility gap can be closed by gathering more information, by making assumptions, by gauging data quality (for data that create inconsistencies), and/or by expanding the story by adding more causal factors and relationships. (See Klein et al., 2019)

Little is gained from trying to work out which of these types of gaps are operating because they overlap so much: information, mental model, plausibility. What matters is that we view curiosity as a motivation to fill a conceptual gap.

Before we proceed, it may be useful to make some clarifications. This paper addresses "epistemic curiosity" -- being curious about events, which is the ***state*** of curiosity. The paper does not examine individual differences in curiosity as a general ***trait***.

Further, this paper is about D-type curiosity, to use Litman's (2019) terminology — curiosity driven by a perceived deprivation of information. In other words, a gap. The paper is not concerned with I-type curiosity, which is about intellectual interest, exploring for the fun of it.

FEATURES OF CURIOSITY

Following Loewenstein, we can see curiosity as superficial, transient, intense, and impulsive.

- Superficial in that curiosity can arise, change focus, and end abruptly. Loewenstein offers the example of a supermarket check out line. While waiting our turn and idly scanning the side racks, we may become intensely curious about the latest news regarding a movie star's marital woes — but this curiosity disappears as soon as we step forward and away from the tabloids.
- Transient in that curiosity reflects attention, and when attention captured by something else, curiosity disappears. See the supermarket tabloid phenomenon described above. That is what is wrong with Claim #1 above, that once curiosity is aroused it persists until it gets satisfied. Sometimes that is the case but at other times the curiosity disappears.
- Intense in that people will work to satisfy their curiosity despite being hungry or thirsty.
- Impulsive in that something can evoke curiosity and stop us in our tracks. We may be sorely tempted to buy that tabloid in the supermarket as long as we are looking at it.

DYNAMICS OF CURIOSITY

Loewenstein has provided evidence showing that curiosity increases as the item of information has more potential to close the gap. Therefore, the smaller the gap, the greater the curiosity — because a smaller gap is so close to being bridged.

That is what is wrong with Claim #2 above, which asserts that the larger the gap the greater the curiosity. Instead, a small gap, with promise of closure, is much more motivating.

When faced with an insight problem, such as the puzzles typically presented during research projects on the nature of insight, the gap is closed by a single piece of information.

Research suggests that if we are very ignorant about a domain, we have little curiosity. That is what is wrong with Claims #2 and #3 above. As the perceived gap is larger, the motivation to explore it decreases. Conversely, as the gap diminishes our curiosity increases. This process is similar to a goal gradient effect in exploring a maze.

BLOCKING CURIOSITY

In reviewing the literature (e.g., Loewenstein, 1994; Kashdan, 2004) and in conversations with experienced training developers, I have compiled a set of different practices that can interfere with curiosity. This list is not exhaustive, and is not based on my own research. However, I hope that readers will find it useful. A warning may be in order — this is a long list and the blockers are fairly common, so readers may become discouraged in reviewing the items.

General Inhibitors

- Overconfidence. We may think we know more than we do, and as a result we experience little or no curiosity to find out more. One type of counter for overconfidence is to have trainees make predictions and provide them with feedback, to reduce their overconfidence.
- Internal pressures such as guilt, fear and anxiety, appear to diminish curiosity.
- External pressures such as threats and punishment, and tangible external rewards all diminish curiosity for specific tasks. For instance, offering people rewards for solving interesting problems seems to reduce their intrinsic motivation to find solutions.

Instructor-initiated inhibitors. These are ways that teachers and instructors can discourage curiosity.

- Ask closed questions — calling for Yes/No answers.
- Focus on procedures.
- Emphasize memorization.
- Claim that there are right answers, which reinforces the importance of memorization.
- Quickly correct students' mistakes, instead of letting students wrestle with the mistakes and make their own discoveries.
- Conduct After-Action Reviews that contain mostly statements and few questions.
- Focus on what the trainee is doing, and not why the trainee is doing it. Focus on the behavior rather than on the thinking process that led to the behavior.
- Overwhelm students with details.
- Introduce too much complexity too early. The problem here is that too much confusion and complexity and ambiguity can inhibit curiosity. The student stops being curious and instead tries to memorize all the material. In addition, by presenting too many open variables and loose ends, instructors can inhibit curiosity.
- Discourage questions. Too often, instructors make it clear that they need to get through all the material and don't have time for questions.
- Discourage class discussions — the same reasoning applies, instructors more interested in getting through the material than promoting discoveries.
- Provide complete explanations to the students and trainees. These complete explanations leave no space for students to engage in self-explanation. (Chi et al., 2014)
- Provide explanations that make students feel stupid. Some instructors use ridicule and as a result, students become afraid to appear dumb.

Provoking curiosity

When I began my investigation into curiosity I had the fear that it was easy to block curiosity but much harder, if not impossible, to promote curiosity. I now believe that this fear was unfounded. My primary objective in preparing this paper was to share the range of ideas for encouraging and using curiosity, particularly in training and educational settings.

General factors that elicit curiosity. In natural settings, these conditions spontaneously arise.

- We encounter events that are novel, complex or ambiguous. We are struck by their surprisingness. Or by an incongruity — some ideas that are incompatible. Or, more generally, by violated expectancies that make us wonder what is going on.
- We are aware of information that is missing or incomplete. Or we may become aware that our understanding is incomplete. Loewenstein provides the example of how we might strain to hear a conversation at the next table in a restaurant. A common experience is when we set out to solve a puzzle, or when we read a mystery story and want to find the identity of the murderer.
- We are having difficulty in predicting the future, and want to learn more.
- We can't remember or find information you once had. For example, the tip-of-the-tongue phenomenon. We know we know the answer, but it isn't coming to us and we are frustrated and motivated to keep thinking about it.
- We want to obtain information someone else has. Another Loewenstein example is when we see someone chuckle when reading a newspaper and for that instant we are dying to know what was in the article.
- We believe in our efficacy. We think, "I can close this gap."
- Pre-existing interests affect our curiosity because the reference point is elevated so the gap is reduced and seems easier to close. That is, our knowledge puts us ahead of the game, making it seem easier to get an answer. And our pre-existing interest in the topic primes the pump of curiosity.
- Contrasts and counterfactuals can evoke curiosity. For example, working with advanced technology we may wonder why did an Artificial Intelligence system do A and not B? Or what would have happened if that value was 250 and not 100.

Instructor-initiated tactics to elicit curiosity. These tactics may have the highest payoff for teachers and trainers. They are methods that can be put into practice.

- Pose a question. This is the simplest method but one of the most powerful. Just asking a question can provoke the curiosity of the learner. Or present questions involving contrasts and counterfactuals and hypotheticals.
- Provide a sequence with an unknown resolution. For instance, who will win an election, or a race? Research suggests that the curiosity will be more intense if we have the students and trainees make their own predictions.
- **Do not encourage students to ask questions. The double asterisk here is intended to highlight this point because it is such a common practice, "Any questions, class?" The suggestion is to pose questions to the students and trainees, such as describing violations of expectancies. This "primes the pump." This recommendation refutes Claim #4 above, that instructors should encourage trainees to ask questions.
- Show students that the gap in their knowledge is manageable. Remember that curiosity is elicited when the gap is smaller, not larger.
- Quiz students/trainees to see how well they understand how a process works. We can use hypotheticals and counterfactuals to help the students and trainees speculate more.
- Wonder what caused a student or trainee to make a mistake. This is the mindset shift mentioned earlier, from a mindset of criticizing trainees for making a mistake to a mindset of wondering what led to the mistake.

CONCLUSION

The set of practices for provoking curiosity show that instructors can readily generate curiosity, refuting Claim #5 ("You can inhibit curiosity but you can't generate it.") Curiosity seems to be an essential capability that can reverse or at least diminish many counter-productive attitudes and behaviors. We can learn to harness curiosity more effectively, and thereby strengthen many of the macrocognitive processes and functions.

ACKNOWLEDGMENTS

I thank Joseph Borders, Devorah E. Klein, John Schmitt, Laura Militello, and Emily Newsome for helpful comments and suggestions, and Adam Zaremsky for his support in producing this paper.

REFERENCES

Chi, M.T.H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49(4)*, 219-243.

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of personality assessment*, *82*(3), 291-305.

Klein, G. (2020, December 3). How to Harness Curiosity? [Blog post]. Retrieved from https://www.psychologytoday.com/us/blog/seeing-what-others-dont/202012/how-harness-curiosity

Klein, G., Hoffman, R.R., and Mueller, S.T. (2019). "The Plausibility Cycle: A Model of Self-Explaining How AI Systems work." Report from Task area 2, DARPA Explainable AI Program.

Litman, J. A. (2019). Curiosity: Nature, dimensionality, and determinants. *The Cambridge handbook of motivation and learning*, 418-442.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, *116*(1), 75-98.

Complexity of field intervention in safety management: a case study in Air Traffic Control

Stéphane DEHARVENGT[a]

[a]*Direction des Services de la Navigation Aérienne (DSNA/DGAC)*

ABSTRACT

This article relates the construction and findings of an intervention focused on Air Traffic Control safety at an airport following an incident. The uncovered root cause is a deviation of communication practice between two air traffic controllers for handling small aircraft traffic on a secondary runway. The intervention consists in coaching and empowering a local project team so that they would work with the air traffic controllers to understand the reasons for the deviation. Guided adaptability is constructed by training air traffic controllers to observe practices in the control tower. The outcome leads to a transformation of local safety management practices. The results are discussed in terms of adaptation of scientific methods to capture normal work and leverage for behavioural changes in attitudes towards deviating from a formal procedure.

KEYWORDS

Adaptation; aviation ; decision making ; situation assessment ; safety management ; field research.

INTRODUCTION

DSNA (Direction des Services de la Navigation Aérienne) is the French Air Navigation Service Provider. The company is certified to deliver air traffic management services at 5 Area Control Centres (so-called CRNA En-Route centres for upper airspace) and more than 80 airports grouped in regional areas (SNA) around major airports, 12 for continental Europe and 3 overseas. The managed airspace represents 1 million km², one of the most important in Europe. It employs about 4000 air traffic controllers to do so. A Safety Management System (SMS) ensures the safety evaluation of any change in the functional system (the combination of technical systems, procedures and human resources) and the analysis of experience feedback (incident analysis), based on the principles of Just Culture to ensure reporting of any safety related issue.

The Air Traffic Management issue

The European Action Plan for the Prevention of Runway Incursions released in 2003 contains the following recommendation: "*improve situational awareness, when practicable, by conducting all communications associated with runway operations on a common frequency. Note - Aerodromes with multiple runways may use a different frequency for each runway*". DSNA dedicated action plan in 2004 includes: "*Ref. 1.2. Crossing runway, management of inter-runway. Unanimously, and taking into account the extreme safety sensitivity of runway operations, all runway operations, […] must be managed by Tower controller on the Tower frequency thus requiring a clearance*". Namely, at any airport, when aircraft are in the air in the vicinity of the airport and when they are moving on the runway(s), they must always be in contact with the TOWER Air traffic controller (TWR ATCO). When they are on the taxiways moving towards the parking or towards their take-off point, they are in contact with the GROUND Air traffic controller (GRD ATCO). Coordination is established between the two ATCOs after an aircraft lands and vacates the runway or when an aircraft waits for penetrating on the runway.

In 2014 an incident led to a landing clearance being delivered on an already occupied secondary runway intersecting the main runway. The error was caught by the pilot of the landing aircraft who initiated a go-around procedure. The investigation showed that the aircraft occupying the runway was in contact with the GRD ATCO. However, the clearance to land was delivered by the TWR ATCO (two different ATCOs on two different frequencies).

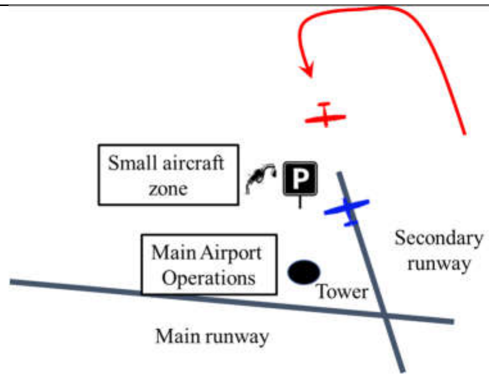


Figure 1: Airport layout description and aircraft positions at the time of the 2014 incident

The local safety committee directed an investigation and decided local corrective actions, but due to the gravity of the potential consequences, the incident was also reviewed during a corporate safety committee. The discussions at that stage uncovered a widely used deviation among the ATCO population at that airport: roughly 1/3 were not following the procedure (TWR ATCO was delegating the task to GRD ATCO to manage traffic taxying on the secondary runway), 1/3 were applying the procedure (TWR ATCO always in charge of traffic in the air and on the runways) and the rest, by group adherence, was following the most vocal group, not following the procedure. Initial explanations pointed to previous local habits (before implementing the "one frequency" procedure), as well as design issues at the airport : the secondary runway does not have a dedicated taxiway, so small aircraft traffic in large numbers have to use the secondary runway to go and park or refuel. That runway is thus being used not only to land or take-off, but also to taxy, contrary to the main runway.

Theoretical framework

This specific incident highlights a local effect of a safety management decision. A European and corporate decision is made and implemented on all large and medium national airports, whatever their size, design or traffic structure: in this case, a procedure is imposed to handle all traffic in the air and on the runway of an airport with a single ATCO on a single frequency. The safety purpose is to decrease the risk of loss of situation awareness. In effect, this corresponds to one of the two distinct modes through which safety is achieved in the safety management literature: either safety management through centralized control, as in our case, or safety management through guided adaptability (Provan, Woods, Dekker, & Rae, 2020). In this paper, the second mode is tested via the empowerment of local stakeholders. The goal is to understand the local practices and decide whether there is sufficient robustness to support a different procedure. The difficulties relate to implementing this kind of approach to safety management of a large organization in a high risk industry and how to learn lessons in terms of methodology to do so.

Aviation is one of the fields where Resilience Engineering (RE) methodologies are being developed and tested, especially for Safety Management Systems (Dijkstra, 2006; Zimmermann, Paries, Amalberti, & Hummerdal, 2011). Indeed, there are many constraints, safety concerns, technological systems, economic and social aspects, as well as the interconnexion of the many stakeholders. Therefore, the system needs to continuously adjust to the many changes or disturbances (man-made, such as modernization of technology or improved greener efficient procedures, or natural, like stormy weather). It does so by dealing with expected and unexpected situations (Hollnagel, Woods, & Leveson, 2006). Two main perspectives can be used to evaluate the conditions for resilience (Paries, Macchi, Valot, & Deharvengt, 2018): one is how control is ensured throughout the system, the second is how adaptations are allowed in the system. The response of the organization in our case was primarily in terms of regaining control of the system – how to eradicate local deviation – but without concern for the local conditions that drive the inevitable trade-offs related to strict implementation of a European wide procedure. However, an organization can manage to investigate the conditions for understanding the adaptation conditions of that procedure. The author, as deputy Safety Director, negotiated a non-judgmental position throughout the process: it is necessary to allow the local stakeholders to make their case and empower them to present substantiated local conditions of control that are said to exist, in order to demonstrate how safety is nether the less maintained.

A Naturalistic Decision Making (NDM) approach is one way to make sense of the intricacies of the decisions made by safety professionals in operational conditions (Klein & Wright, 2016). Such an in-depth analysis taking into account field settings is relevant because of the performance issue (Klein, 2008): Air Traffic Management (ATM) is a complex system where ATCOs pursue different goals: speed up the flow of aircraft, reduce the environmental impact of aviation (in the case of an airport this relates to the noise issue for surrounding populations). They also "help" pilots by making their life (or more exactly their flight) easier, that is shorten the trajectories to save fuel or allow them levels of flight for more efficient fuel consumption and aircraft performance. At the same time this system is very much constrained: by procedures like rules of the air, by flight routes like standard departure or arrival, and by various coordination mechanisms between air control centres or airspace sectors. This field setting corresponds to those "*vaguely defined goals in the face of uncertainty, time pressure, high stakes, team and organizational constraints, shifting conditions, and action feedback loops that enable people*

to manage disturbances while trying to diagnose them" (Orasanu & Connolly, 1993). Decisions made by ATCOs concerning specific procedures must be contextualised to evaluate how safe they are for the system.
Taken together, RE and NDM offer a good and confluence ground to analyse complex local operational decisions in the framework of organisational safety management.
Finally, the case demonstrates continuity from Safety-I to Safety-II management practices (Hollnagel, Leonhardt, Licu, & Shorrock, 2013). The analysis of the incident process is illustrative of the 'centralized control' mode of safety management (Provan et al., 2020): an incident leading to local safety committee review first, then a corporate safety committee review, where corrective actions to remedy the identified causes are supposed to be taken. This classic Safety-I approach, report - analyse - correct, is needed to uncover the issue (deviation of procedure). Because those interactions occur and thanks to a mature safety culture across the organisation, the identification of the deviation permits a novel approach to investigate the existing local adaptation processes. In this sense, the Corporate Safety Management initiates the Safety-II process that is described below.

METHOD

The proposed field intervention

In agreement with the Corporate Operations Director and the Head of Unit, the author, deputy Corporate Safety Director and one of his expert in safety risk assessment, both trained in human factors, directed a project to understand the underlying rationale of the situation. Based on an action plan managed by the Unit, the project consists in a 2 years coaching activity to the benefit of a local ad-hoc team : a project lead and middle management, the local investigator, several ATCOs working part time in the office or as team resource management trainer. The project phases are to analyse past safety incidents, train ATCOs to perform over the shoulder observations in the control tower, develop an analysis grid for observations to be used by the trained ATCOs and analyse the collected data to trigger a transformation of local practices. The following figure illustrates each role and activity in the structured intervention.

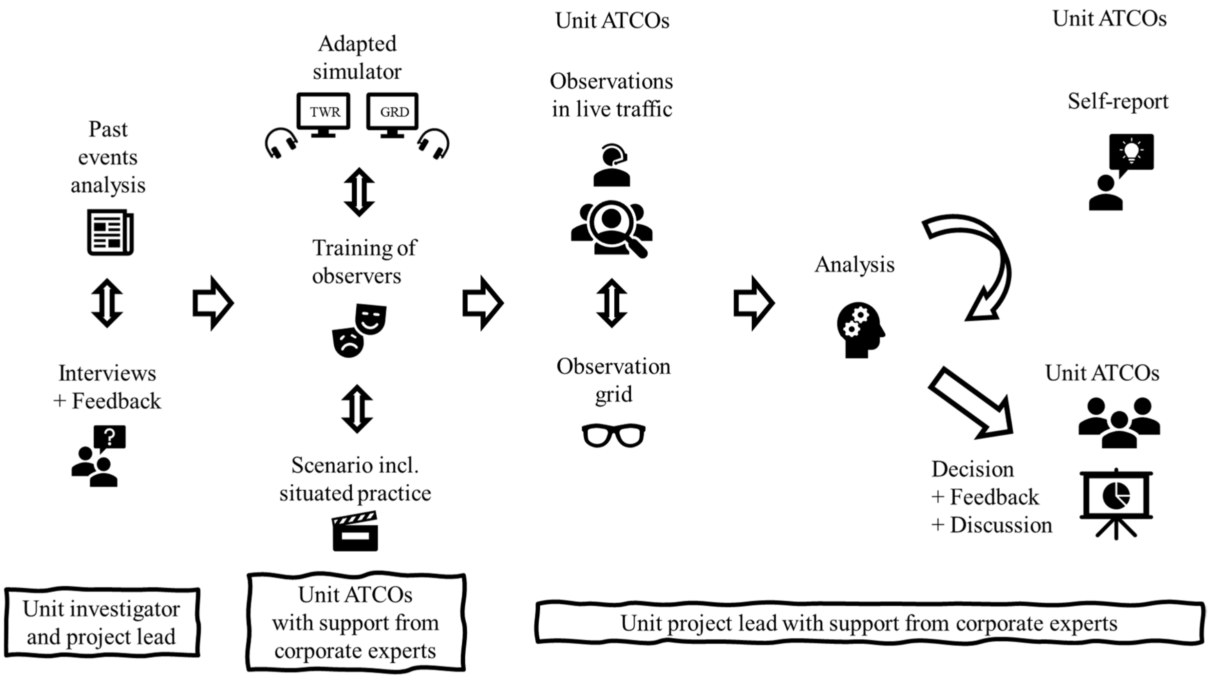


Figure 2: Workflow of the field intervention illustrating roles, methods and outputs

Data collection specifics

Data collection follows the pattern of a safety risk assessment method specifically designed to take into account operational changes that impact ATCOs practices ("volet opérateurs de l'étude de sécurité" – user centred aspects of safety risk analysis of a change – approved by the National Competent Authority). The method is built on past risk analysis of major changes (e.g. transition from paper to electronic stripping in Area Control Centre environment). It is already used successfully for several operational procedure change (e.g. an important change at a major airport consisting in implementing new categories for wake turbulence separations). Over the shoulder observations in normal operational conditions (or normal observations monitoring – NOM (CANSO, 2021)) are a crucial aspect of the method. The intention is to capture existing practices that may be impacted by the change and evaluate how the change will be adopted by the ATCOs. Those observations are usually carried out in simulator, sometimes in the control room or tower, and mostly by trained ergonomist.

The context is slightly different in this case, since the "change", i.e. the non-implementation of the procedure, is already "somehow" or partly in place. The method provides a structured approach to the evaluation of the robustness of the locally developed operational practice.

Past incidents are collected to evaluate the potential contribution of the practice in terms of safety; ATCO supervisors (in charge of the control tower management and strategic decisions during the shift) are also asked to take note when the frequency delegation occur. Several interviews, meetings and visits are performed to capture additional contextual information and frame the implementation of the method. Then some ATCOs are trained to perform periods of observations in the control tower to collect data in a structured way. After this main series of data collection and analysis, a second limited phase is initiated by providing a notebook in the control tower to allow individual ATCOs to freely write their perspective on the instances of frequency delegation and share their related stories.

The team of two experts from Corporate Safety Department always acts in support of the unit project lead and investigator, not performing the data gathering or analysis themselves but guiding the project team and answering questions.

Normal operations monitoring

Performing observations in real operational conditions is an essential aspect of the process. Unit ATCO peers are selected to be trained as observers. The length of the observation periods is difficult to determine with sufficient advance notice, that is when traffic conditions is opportune for some ATCOs to use the delegation of frequency. Therefore, the availability of the observers at rather short notice is needed. Those selected work part time in the office at the bottom of the control tower, making it easier for them to be available. As local ATCOs observe their peers, the obvious bias of not knowing the local operational context and local practices, important for air traffic control, can be alleviated. Otherwise specific local issues unrelated to the topic observed are captured, overwhelming the observer, and drowning useful data in "normal noise". A structured observation grid serves to compensate for local rationality bias (observer projecting his/her expertise and understanding on the situation and actions observed). It also helps with cross-observer standardisation of collected data.

This grid is constructed based on the different feedback sources gathered: when to observe, what to observe, who initiated the action of delegation, what is the operational context (including weather conditions, important when dealing with small aircraft traffic). The main objective is to evaluate how the coordination between TWR ATCO and GRD ACTO is organised when the delegation of frequency is used.

For the training, the unit simulator is adapted to emulate both positions (TWR and GRD). A short scenario is created to allow the observers to be trained to introduce themselves to their colleagues and what they are going to do, so as to gain acceptance from the observed ATCOs in the tower. They learn how to fill the observation grid during the scenario while capturing factual elements. They also conduct an interview of the observed ATCOs after the sequence to validate what they understood. Finally, they collect additional comments from the perspective of the observed ATCOs to facilitate interpretation of data. The training sessions are conducted by the Corporate Safety Department team (a trained ergonomist and an experienced human factors expert). The training lasts for 2 days and 4 ATCOs are trained: ½ day for presenting the method, 1 day for playing the scenario to the different trainees at the simulator, ½ day for debriefing). 2 additional ATCOs are needed to play the scenarios in the simulator.

Performing observations is a specific social process, even more so in an air traffic control tower. Social aspects of working in a team of ATCOs are highlighted in several ethnographic studies (Poirot-Delpech, 1995; Potthast, 2008). Accepting observers, even peers, in a close setting when safety is at stakes, when operational decisions are made, especially when the topic to be observed deals with differing practices, needs to be negotiated with care and openly communicated. It also serves to prepare the acceptance of the products resulting from the intervention: those data are collected by their peers (licensed ATCOs with current unit endorsement) using a structured method in their own operational working environment with normal live traffic, without direct corporate or middle management involvement.

RESULTS

Data analysis

The first incidents analysis and interviews highlight some of the reasons for the deviation in practices. Since the frequency delegation was used before the new procedure was made mandatory, it is considered as an "ancestral delegation" (verbatim). Local aircraft traffic and airport design features also appear to support its continuation. Indeed, small aircraft frequently use the secondary runway in order to refuel or to park (flying club), because there is no taxiway to circumvent using the runway for this purpose. From the TWR ATCO perspective, delegating the surveillance of the aircraft on the secondary runway makes sense because that runway is mostly in his back (see the location of the control tower in Figure 1). Thus, he can keep monitoring other traffic on the main runway in front of him rather than having his attention distracted from less relevant aircraft (verbatim: "monitoring small aircraft taxying for 3 minutes - inapplicable"). "Appropriate" coordination is said to prevent the TWR ATCO to

allow landing on the secondary runway while another aircraft is already using it to go to the parking or refuel under the monitoring of the GRD ATCO by delegation.
Additionally, crossing runways at the intersection of main and secondary runway by small aircraft is usually done on the same frequency. Indeed, this only requires a very short time, while transferring the aircraft to the TWR, allowing the crossing, then transferring back to the GRD would take too much time. A nationwide analysis reveals that for other airports with intersecting runways, keeping the aircraft with GRD is also common practice in order to speed up traffic and not bother the TWR ATCO. The main difference, however, is that in most cases the other runway is not active at those other airports, contrary to what is the case here.
During a 3 months period, the 4 observers conduct several hours of observations, leading to 35 instances of observed frequency delegations. Debrief interviews are also conducted. However, no clear picture emerges from the data collected. Especially, and despite the claims of those supporting the delegation, no clear obvious coordination method is identified. This remains the main argument of the supporters of the delegation in that the "coordination" ensures the robustness of this practice in view of the potential risks (e.g. forgetting an aircraft was using the secondary runway when another is cleared to land on the same runway in the back of the TWR ATCO). The various operational contexts where the delegation is used also do not provide clear indications for developing a shared best practice.
The absence of a clear risk mitigation solution is a trigger for allowing a complementary experiment. The unit ATCOs are debriefed about the findings of the observations period and informed that further substantiation is needed if the practice is to be continued. A defined period when the delegation would be officially authorized is proposed in exchange for ATCOs to provide self-reports on their usage of the delegation with operational context and detailed information. However, very few reports are subsequently made, and it is decided jointly by Corporate Safety and the Head of Unit to terminate the experiment.

Behavioural changes

Despite (or thanks to) the lack of result to support the delegation of frequency, an observable change is witnessed by the unit management. Via an informal survey among the ATCOs, it is recorded that the 1/3 part of the population that was following the most vocal supporters of the delegation, now is following the official procedure, keeping all aircraft in frequency with the TWR ATCO.

DISCUSSION

Implications for Theory

Safety defined as a "dynamic non-event" (Weick, 1987) drives the focus on the dynamics of practice and how it manages variability in a dynamic environment. The example in this article highlights how social constructs, and their evolution, influence the understanding of what constitute safety in the eyes of the professionals and how to best maintain it.
The method used calls for a specific usage of structured observation methods. This method is the trademark of ergonomists affiliated to the French Speaking Ergonomist Society (definition of ergonomics by the International Ergonomics Association in August 2000). The purpose is the analysis of actual work situations with a view to transform the work situation. Here, the intervention is not performed from beginning to end by a trained ergonomist. However, the whole process is supervised by one. Although not formally recorded and demonstrated, the work situation is actually changed thanks to the commitment and involvement of local stakeholders as well as to the leverage obtained from performing observations and disseminating their outcome. NDM is at the heart of the process by investigating the decisions made by experts to use or not a specific professional practice depending on the operational situation. No specific coordination mechanism exists between the two ATCOs: this product of observed normal operational situations, is the key driver for a change in practices for an important number of professionals in this unit.
It appears that the overall resilience of the unit is improved by the change of view regarding the practice. The perspective on the trade-offs that were viewed as necessary (maintaining focus on the main runway vs. devoting time to monitor the secondary runway) shifted for a number of ATCOs: keeping a "complete traffic picture" as part of their cognitive mental process (including small aircraft only using the secondary runway as a taxiway) seems to be more important than having a partial image by delegating a small part of the responsibility to another colleague, without the proper mechanism that would ensure a safe and secure collaborative work. This seems to validate the initial rationale for the procedure, i.e. situation awareness.

Implications for Practice

This article presents a practical implementation of a Safety-II method. The identified deviation from procedure (typical Safety-I) is scrutinized to understand the adaptability conditions and their acceptability for safety, rather than be met by further pressure for compliance. The responsibility to investigate is decentralized from Corporate headquarters to the unit. The local team is empowered by facilitating the implementation of the method, including the development of what needs to be observed as well as training ATCOs as peer observers. The analysis of work-as-observed is thus guided (Provan et al., 2020). The company management system then operates in dual modes on demand.

The intervention also benefitted from years of over the shoulder observation by researchers involved in Air Traffic Control. Observations techniques including bias and ways to overcome those are well known. However, they are not very well understood or even feared by managers or by ATCOs for different reasons: costs and resources, subjectivity, big brother effect, complexity, open to manipulation by those observed, … However, coupled with the appropriate subject matter expertise, as well as other well-known techniques like interviews or questionnaires, Normal Operations Monitoring (NOM) can provide key insights into the trade-offs of normal work under dynamic circumstances. Several Air Navigation Service Providers have now implemented this method under various schemes and purposes (CANSO, 2021). The knowledge and insights gained are reasons to promote the method as another standard in the Safety Management Systems' toolbox and contribute to the evolution towards a more resilient approach to Safety Management.

CONCLUSIONS

The article presents the case of a risk inducing practice at an airport that was not shared uniformly among the group of ATCOs. The practice contradicts the formal operational procedure, recognised internationally as a standard for safety. An approach involving NDM method to highlight how trade-offs are accepted without secure mitigation means allows a change in practice towards a more resilient management of the constraints of the airport layout and traffic.

Although the results lack robustness, the benefits from a Corporate or local Unit Management perspective are important since the change is brought without resistance and undue internal discussions. Verifying that a drift back into the old habits has not reoccurred is however needed.

The feasibility of such an approach is also demonstrated together with the caveats that a minimal strong scientific leadership is a must. Also, the time span to produce results is long and the outcome is in no way predictable from the onset. This example indicates that changing the mindset of most safety managers may be needed. The accepted practice is that specific targeted actions towards identified risks must be put in place and verified in order to achieve a certain state of safety in terms of conformity to identified procedures. Rather, monitoring systems for capturing normal practices and drift or adaptations, are proven useful to guide them into unchartered territory. It constitutes a major step change for Safety as well as Board Managers. That's a trade-off worth looking into since it offers an opportunity for a joint undertaking between managers and field matter operational experts.

REFERENCES

CANSO. (2021). *Normal Operations Monitoring Toolbox. A systemic approach to everyday safety*. Retrieved from https://canso.org/publication/canso-normal-operations-monitoring-toolbox/

Dijkstra, Arthur. (2006). *Resilience Engineering and Safety Management Systems in aviation.* Paper presented at the 2nd symposium on Resilience Engineering Juan-Les-Pins, France.

Hollnagel, Erik, Leonhardt, Jörg, Licu, Tony, & Shorrock, Steven T. (2013). *From Safety-I to Safety-II: A White Paper*. Retrieved from http://www.skybrary.aero/bookshelf/books/2437.pdf

Hollnagel, Erik, Woods, David D., & Leveson, Nancy G. (2006). *Resilience Engineering - Concepts and Precepts*: Ashgate.

Klein, Gary. (2008). Naturalistic Decision Making. *Human Factors, 50*(3), 456-460. doi:10.1518/001872008x288385

Klein, Gary, & Wright, Corinne. (2016). Macrocognition: From Theory to Toolbox. *Perspective*. http://dx.doi.org/10.3389/fpsyg.2016.00054 doi:10.3389/fpsyg.2016.00054

Orasanu, Judith, & Connolly, T. (1993). The reinvention of decision making. In G Klein, J Orasanu, R Calderwwod, & C E Zsambok (Eds.), *Decision making in action: Models and Methods* (pp. 3-20). Norwood NJ: Ablex.

Paries, Jean, Macchi, Luigi, Valot, Claude, & Deharvengt, Stéphane. (2018). Comparing HROs and RE in the light of safety management systems. *Safety Science*. doi:https://doi.org/10.1016/j.ssci.2018.02.026

Poirot-Delpech, Sophie. (1995). *Biographie du CAUTRA. Naissance et développement d'un système d'informations pour la circulation aérienne.* (Doctor), Université Paris 1 - Panthéon Sorbonne, Paris.

Potthast, Jörg. (2008). Ethnography of a Paper Strip: The Production of Air Safety. *Science, Technology & Innovation Studies, 4*(1).

Provan, David J., Woods, David D., Dekker, Sidney W. A., & Rae, Andrew J. (2020). Safety II professionals: How resilience engineering can transform safety practice. *Reliability Engineering & System Safety, 195*, 106740. doi:https://doi.org/10.1016/j.ress.2019.106740

Weick, Karl E. (1987). Organizational Culture as a Source of High Reliability. *California Management Review, 29*(2), 112-127. doi:10.2307/41165243

Zimmermann, K., Paries, Jean, Amalberti, René, & Hummerdal, H. D. (2011). Is the aviation industry ready for resilience? Mapping human factors assumptions across the aviation sector. In *Resilience Engineering in Practice: A Guidebook* (pp. 257-274): Ashgate Publishing Ltd.

How Nursing Care Stations and Care Facilities in Japan Coped with COVID-19 Pandemic: A Survey of Challenges, Preparations, and Responses

Taro KANNO[a], Mizuki UMEMOTO[a], Chie ISHIDA[b], Aya IGUCHI[b],
Jun SATO[c], Risa OKADA[d], and Takayuki KANESAKA[d]
[a]*The University of Tokyo*
[b]*Japan Red Cross College of Nursing*
[c]*Tokyo Healthcare University*
[d]*Carepro Inc.*

ABSTRACT

This study presents a questionnaire survey, designed by a resource-centric framework, about how Japanese home-visit nursing stations and nursing care facilities prepared and coped with the COVID-19 pandemic. We analyzed the answers using a qualitative analysis to develop categories of good and typical practices to cope with pandemic situations. Based on the analysis results and discussions, we developed five categories of preparedness and countermeasures for resource management: protection, utilization, stockpiling, procuring, and repairs. Furthermore, we provided the categories with practice examples for human resources, which can be used as a reference when preparing new business continuity plans (BCPs).

KEYWORDS

Problem-solving; crisis management; business continuity plan; COVID-19; healthcare

INTRODUCTION

The importance of developing a practical business continuity plan (BCP) has been widely acknowledged in many sectors of society, particularly in the medical and healthcare industries that are expected to continue the business and service to save injured and vulnerable people during disasters. For example, in 2019, the Japanese Ministry of Health, Labour and Welfare (MHLW) revised the notice and required disaster base hospitals (DBHs) to formulate BCPs (MHLW, 2019a). In response to the COVID-19 pandemic, the MHLW recently announced that they also demand nursing care facilities to prepare BCPs in 2020 (MHLW, 2020). A survey on the readiness of BCPs in DBHs in 2019 (MHLW, 2019b) revealed that almost all DBHs have finished formulating BCPs in these two years. However, their workability and effectiveness are doubtful because there are no good and easy-to-use guidelines for developing BCPs, and there is no established method to evaluate them. These problems are more problematic for nursing care facilities because they do not have sufficient knowledge and experience for BCP development. In addition, because their business size is much smaller than that of the DBHs, they typically do not have enough personnel and money to be assigned for BCP preparation. To solve these difficulties in preparing and evaluating BCPs, we proposed a resource-centric BCP framework that provides categories of the things to be considered in BCP on four types of resources for business continuity and recovery (Kanno, Ishida, Kanesaka, Okada, Kawazoe, Tae, & Sato, 2019).

This study introduces a survey, designed by a resource-centric framework, about how Japanese home-visit nursing stations and nursing care facilities prepared and coped with the COVID-19 pandemic. In addition, we provide a preliminary report of the survey results. In the next section, we explain a resource-centric BCP framework and its recent modifications. Section 3 presents an overview of the survey and the questionnaire design based on a framework, which explores good and typical practices for each type of resource for business continuity and recovery. Section 4 provides an overview of the survey results and a detailed analysis of the human resources. The discussion and conclusions are provided in Sections 5 and 6, respectively.

RESOURCE-CENTRIC BCP FRAMEWORK

In our previous study (Kanno et al., 2019), we proposed a framework for the simple design and evaluation of BCPs, which is characterized by the concept that resource management is the essence of disaster management, including preparedness, response, and recovery. A conceptual formulation of a resource-centric BCP is presented in Equation (1). This framework defines four types of essential resources and three purposes for resource management. The details of these categories are described in the following subsections.

$$\text{BCP contents} = Resources \times (Prevent\ loss + Prevent\ decrease + Promote\ Increase) \quad (1)$$

Four Basic Types of Resources Plus One

In ISO 22301:2019, the resource is defined as *"all assets (including plant and equipment), people, skills, technology, premises, and supplies and information (whether electronic or not) that an organization has to have available to use, when needed, in order to operate and meet its objective."* While this provides several resource types, they are not comprehensive, and their definitions are not clear. In a previous study, we defined four types of essential resources required for business continuity and recovery: *Hito* (human resources), *Mono* (physical resources), *Kane* (monetary resources), and *Jyo-ho* (data and information). These categories are still not comprehensive, but they are commonly known as management resources or firm resources, familiar and easily understood by BCP nonexperts. Additionally, we listed a new resource type, *Shinrai* (credibility or reliance), which is crucial when communicating with service users for service-level degradation and acquiring new users during and after a disaster.

The resources of the first three types are either consumable or non-consumable. While consumable resources disappear or cannot be used again, nonconsumable resources are just occupied and released after they are used. The other two are nonconsumable but are not occupied. Another aspect to be considered is quality, which is further categorized into attributes and properties. The attribute does not change or deteriorate when used, while the property is. For example, knowledge and skill are regarded as human attributes, while vitality/fatigue is a property.

Three Basic Purposes for Resource Management

The ultimate goal of the resource-centric BCP is to keep all the resources sufficient to satisfy demands. However, we cannot realistically expect it at the time of a disaster; rather, it is not a disaster if the goal is achieved. Therefore, what we need to consider in BCP is to prepare for resource shortages. We categorized three primary purposes for resource management to cope with resource shortage: (1) prevent loss, (2) prevent decrease, and (3) promote increase. Loss prevention refers to the protection of resources from loss, damage, or deterioration caused by disaster hazards. The prevention of decreasing means utilizing available resources to continue business as much as possible. The increasing refers to increase the quantity and quality of resources to maintain business continuity and finally fully recover to satisfy the demand.

QUESTIONNAIRE SURVEY

Questionnaire Design

We conducted a nationwide questionnaire survey to find practical preparations and responses in nursing care facilities to cope with the COVID-19 pandemic. Questionnaires were designed based on a resource-centric BCP framework to collect data effectively and comprehensively. In the main part of the questionnaire, we asked potential and actual problems that will lead to the shortage of each type of resource and preparation and actual responses taken to achieve the three primary purposes of resource management. An excerpt of the questionnaire, questions about human resources, is presented in Table 1. Q.1 asks questions asking about problems. Q.2, Q.3, and Q.4 correspond to the questions asking about prevention of resource loss, prevention of resource decrease, and promotion of resource increase. The multiple-choice items were made based on discussions with co-authors, which we assumed as typical possible answers.

Table 1. Questions on human resources (staff and employees)

Q.1-a	What kind of problems leading to human resources shortage occurred?/did you anticipate? Select all applicable choices: 1) Infection 2) Suspected infection 3) Close contact 4) Increase in tasks because of the pandemic 5) Decease in productivity because of telework 6) Leave of absence or turnover
Q.1-b	Except for the above choices, what kind of problems leading to human resources shortage occurred?/did you anticipate?
Q.2-a	What kind of countermeasures or responses did you take/prepare to prevent decreasing human resources? Select all applicable choices: 1) Infection prevention 2) Prevention of the spread of infection 3) Teleworking/online meeting 4) Mental care
Q.2-b	Except for the above choices, what kind of countermeasures or responses did you take/prepare to prevent decreasing human resources?
Q.3-a	What kind of countermeasures or responses did you take/prepare to continue the business under human resources shortage? Select all applicable choices: 1) Reduction of service 2) Shortening of working hour 3) Prioritizing services

	4) Adjusting shifts/staff assignment 5) Delegation of authority and responsibility 6) Working harder
Q.3-b	Except for the above choices, what kind of countermeasures or responses did you take/prepare to continue the business under human resources shortage?
Q.4-a	What kind of countermeasures or responses did you take/prepare to increase human resources? Select all applicable choices: 1) New hiring 2) Support retired staff to return to work 3) Ask for external support 4) Wait for the situation to settle down
Q.4-b	Except for the above choices, what kind of countermeasures or responses did you take/prepare to increase human resources?

Survey

We conducted a survey using the questionnaire described in the previous subsection. The survey period was from December 5, 2020, to January 25, 2021. Figure 1 shows the number of polymerase chain reactions (PCRs) positive daily in Japan, which suggests that this survey was conducted during the third wave of COVID-19 pandemic.

We selected 1,000 home-visit nursing stations nationwide from all the National Association for Visiting Nurse Service members using stratified random sampling and sent the questionnaire by postal mail. We also prepared an online questionnaire accessible from a QR code printed on the questionnaire; the respondents could answer the questionnaire either paper or online. The primary purpose of this survey was to collect good practices to cope with infectious diseases in home-visit nursing stations and nursing care facilities, and the secondary goal was to develop a prototype sample of the resource-centric BCP based on the survey results. This survey was approved by the Ethical Review Committee of the School of Engineering at the University of Tokyo.

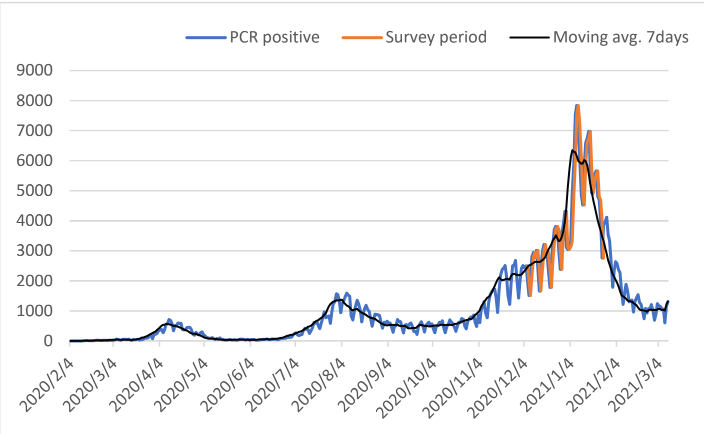


Figure 1. The number of PCR positive daily in Japan[1]

RESULTS

Overview

We obtained a total of 239 responses (response rate: 23.9%). Among them, 159 (66.5%) answered the paper questionnaire, and 80 (33.5%) responded online. Figure 2 shows the rate of the respondents' business size by the number of full-time nurses, which suggests that most of the home-nursing stations' business sizes are relatively small. Figure 3 shows other services and functions provided by the respondents' home-visit nursing station, which indicates that home-visit nursing stations offer various types of nursing and healthcare services. Therefore, the survey results and their implications are expected to be valid to some extent for other kinds of nursing care facilities.

In the questionnaire, we asked about problems and countermeasures on all five primary resources: human, physical, money, information, and credibility. However, in this study, we focus on human resources and report the survey and analysis results.

[1] The original data is available at https://www.mhlw.go.jp/stf/covid-19/open-data.html

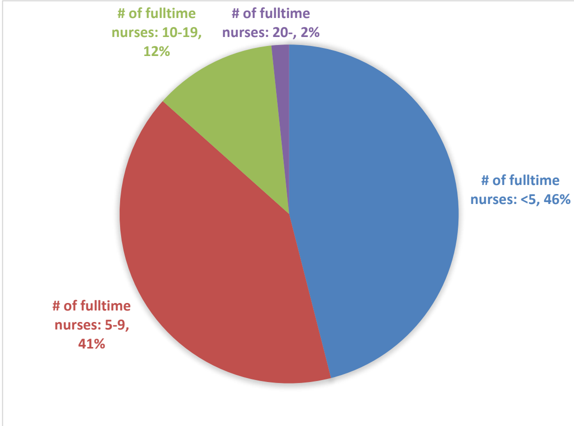


Figure 2. Business size of the respondents' facilities

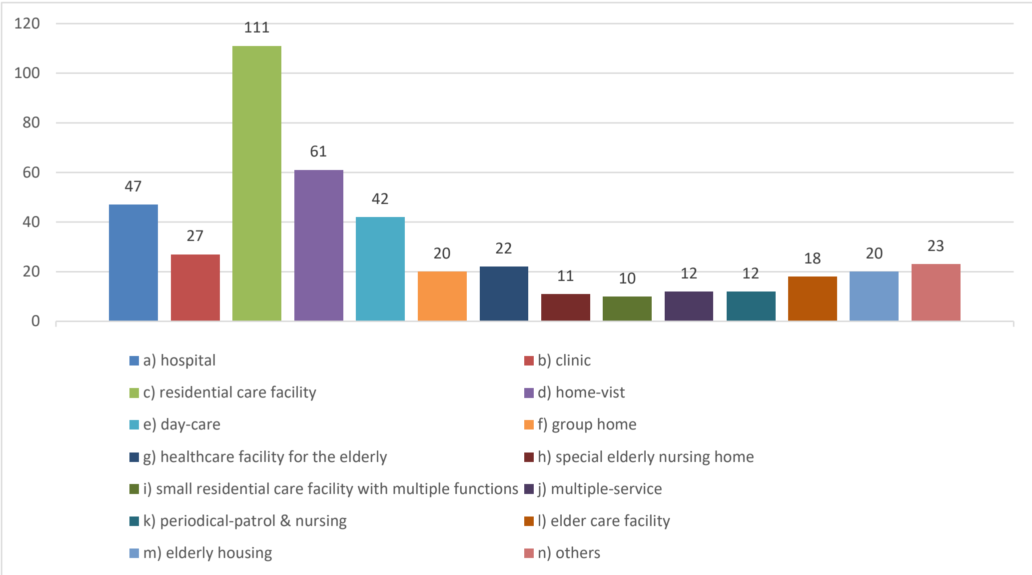


Figure 3. Services and functions by the respondents' home-visit nursing station

Human Resources (Staff): Quantitative Results

Figures 4 to 7 show the answers to the first sub-question (a) for each question (Q.1–Q.4). These questions ask about the actual and potential problems/responses; thus, these results do not mean that these problems or responses actually occurred or were taken.

From Figure 4, the potential or actual problems directly related to the pandemic, such as infection, suspected infection, and close contact, had high percentages. On the other hand, most of the respondents did not find a productivity decrease by teleworking. One possible reason is that they noticed that they could perform many essential tasks, such as meetings and conferences, effectively online because they had already experienced the first and second waves. In addition, the primary task, home visit, cannot be done by telework; therefore, their main business could not be susceptible to telework introduction.

Figure 5 indicates that almost all the respondents took countermeasures or preparations for infection. On the other hand, because the spread of infection differs greatly from region, a sense of crisis and preparations could differ from region to region. Although further analysis is necessary, the answers of approximately 50% might reflect such a background. However, it is notable that the respondents perceived the importance of mental care.

Figure 6 shows that prioritization or triage was considered as the fundamental countermeasures to resource shortage. In practice, prioritization is also necessary for the reduction of service and adjusting shifts. The result also shows that the delegation of authority and responsibility were not considered necessary; this could be because, unlike other natural disasters, it is unlikely for managers or executive staff to be heavily injured by hazards or disconnected because of communication failures.

Figure 7 indicates that it is relatively common for respondents to ask for external support to cope with human resource shortages. The answer for Q.4-b also revealed that it is familiar to ask for support not only from the local or national government but also from neighboring home-visit nursing companies.

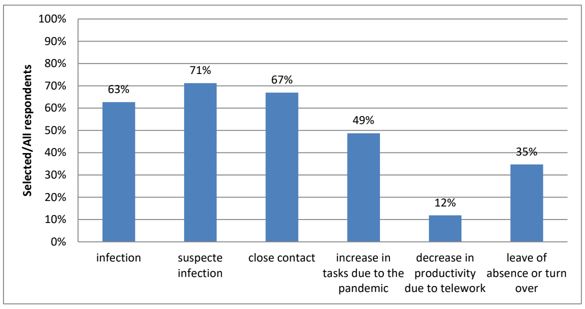


Figure 4. Answers to Q.1-a: Problems

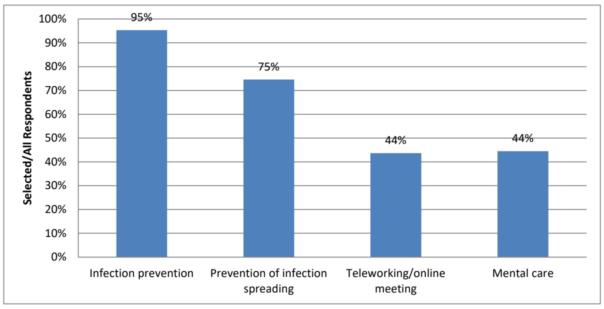


Figure 5. Answers to Q.2-a: Prevent loss

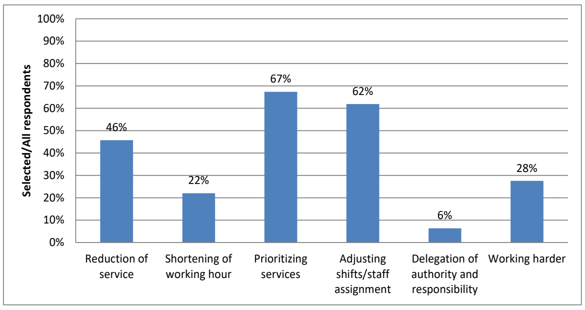


Figure 6. Answers to Q.3-a: Prevent decrease

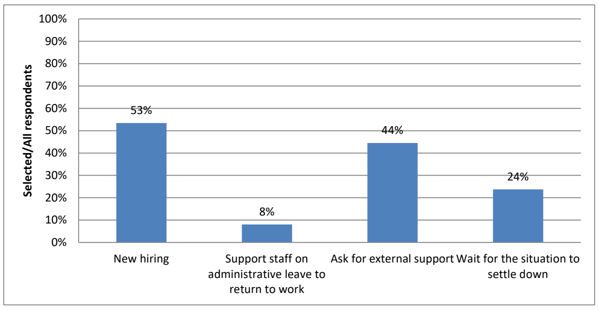


Figure 6. Answers to Q.4-a: Promote increase

Human Resources (Staff): Qualitative Analysis Results

We intended to prepare the second sub-question (b) to collect excellent practices to cope with the COVID-19 pandemic. Because these are open questions, we applied a qualitative content analysis method to summarize the answers. Table 2 shows the number of answers for each question.

Table 2. The number of answers for each sub-question (b)

Q.1-b	**56**
Q.2-b	**48**
Q.3-b	**42**
Q.4-b	**31**

Preparations and Countermeasures

We assumed three purposes for resource management—preventing loss, preventing decrease, and promoting increase—as the top categories. We further developed subcategories for preparations and countermeasures from the answers to Q.2-b, Q.3-b, and Q.4-b from the resource-centric perspective.

Table 3 shows the categories developed through qualitative analysis. The second-level categories are general and common to all types of resources, and the third level is specific to human resources and the context of infectious diseases. We further developed two subcategories for utilization: business reduction and resource extension, which suggests both "degradation" and "extension" were necessary and adopted in practice for business continuity and recovery.

Table 3. Category of preparations and countermeasures (human resources, COVID-19)

Three Purposes	**Preparations and Countermeasures**			**Examples**
Prevent Loss	Protection		Infection prevention	Staggered commuting, go straight client's and right back home, disinfect, limit contact duration, etc.
Prevent Decrease	Utilization	(Business) Degradation	Outsourcing	Within a company, other companies, etc.
			Partial suspending	Decline new contracts, limit the frequency of home-visit, etc.
			Change of business content	Home-visit to teleconsultation, etc.
		Extension	Role re-assignment	Assign multiple roles, increase visits of a single staff, etc.
			Hard-working	Give up a holiday, extend working-hour, etc.,
Promote Increase	(Re-) Sourcing		Hiring	New hiring, convert part-time workers to regular, etc.
			Accepting support	From within a company, from other companies, etc.

Management for Preparations and Countermeasures

As shown in Table 4, we also developed categories of managerial issues included in the answers to Q.2-b, Q.3-b, and Q.4-b. The managerial issues here mean those required monitoring, response, and decision-making to cope with changing situations and trade-offs. In other words, these activities are necessary to efficiently and effectively perform the preparations and countermeasures in Table 3.

Table 4. Category of management (human resources, COVID-19)

Three Purposes	Management for Preparations and Countermeasures			Examples
Prevent Loss	Protection		Infection control	Information sharing
				Health monitoring and PCR test
				Quarantine
				Education and instruction
Prevent Decrease	Utilization	(Business) Degradation	Prioritization	Triage of clients and operations
		Extension	Labor control	Reward adjustment
				Shift adjustment
				Mental care
Promote Increase	(Re-) Sourcing		Human resource management	Clarifying necessary human resources
				Keeping contact information on potential candidates
				Appropriate human resource allocation

DISCUSSION

We confirmed that all the answers and developed categories were categorized under the top three categories, suggesting that the proposed resource-centric framework is valid.

Tables 3 and 4 summarize the answers to the questionnaire survey; however, at the same time, we can say that they are the lists of the things that should be considered in BCPs. In other words, we can say that the questionnaire survey based on the framework is a method to develop a practical BCP efficiently using collective intelligence. On the other hand, as shown in Figures 2 and 3, the business forms and sizes of home-visit nursing stations are diverse. If we further collect and analyze data by different business forms and sizes, we can expect to develop BCPs that are best suited to them.

Restructuring BCP Contents

Based on the analysis and discussions with co-authors, we aimed to restructure BCP contents to simplify its design. Figure 4 shows the structure that considers the disaster timeline, and Table 5 provides plans that constitute a resource-centric BCP. The number allotted to the items in Figure 7 corresponds to that in Table 5. We have not yet analyzed the answers to the questions on other resources; thus, we have not developed categories for these resources. However, stockpiling and repairs are common preparedness and countermeasure for non-human resources such as physical materials. This restructuring or recategorization is not special or innovative. However, because it focuses only on resources and explains what we need to consider in BCPs in terms of actions for resources, we can expect that it is easier even for BCP nonexperts to understand the concept and prepare BCPs.

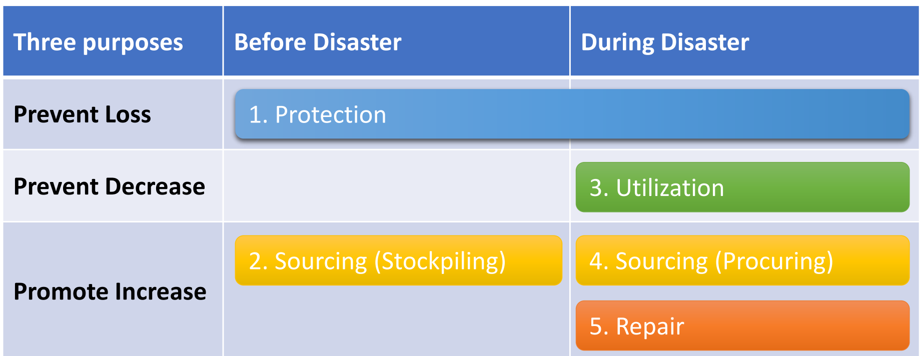


Figure 7. BCP parts along the disaster timeline

Table 5. Resource-Centric BCPs

1	**Resource Protection Plan**	This plan formulates preparedness and countermeasures to protect resources from disaster hazards to prevent resource loss.
2	**Resource Stockpiling Plan**	This plan formulates an appropriate amount of stockpiling to prevent resource shortage.
3	**Resource Utilization Plan**	This plan formulates how to use available resources effectively to continue business when resources are in shortage.
4	**Resource Procuring Plan**	This plan formulates how to procure resources quickly to alleviate and resolve resource shortages.
5	**Resource Repair Plan**	This plan formulates how to repair damaged or broken resources quickly to alleviate and resolve resource shortages.
6	**Resource Management Plan**	This plan formulates how to implement the above five plans and the entire resource management effectively.

CONCLUSION

This study presented a questionnaire survey, designed by the resource-centric framework, about how Japanese home-visit nursing stations and nursing care facilities prepared and coped with the COVID-19 pandemic. The answers contained many excellent and typical practices in home-visit nursing stations and nursing care facilities against the COVID-19 pandemic. We developed categories of preparedness, countermeasures, and management of human resources against infectious diseases using a qualitative analysis of the answers obtained from the open questions in the questionnaire. Based on the analysis results, we offered five basic actions for resource-centric BCPs: protection, utilization, stockpiling, procuring, and repairs. The well-organized categories with these practice examples can be utilized as a good reference when preparing BCPs for infectious diseases.

In the next step, we will analyze the rest of the results on other types of resources, such as human and physical resources, including materials and tools, money, information, and credibility. Subsequently, we will complete the development of categories for preparedness and countermeasures for resource-centric BCPs. In addition, we will extend this survey and analysis to a method for developing resource-centric BCPs by collective intelligence.

ACKNOWLEDGMENTS

We sincerely appreciate the great help and cooperation offered by the respondents, even under such a difficult period.

This work was partly supported by JST-Mirai Program Grant No. JP20348594, Japan.

REFERENCES

Japanese Ministry of Labour Health and Welfare (2019a). Requirements for disasater base hospitals, Notice by the director general of Health Policy Bereau, 0331-33 (in Japanese): https://www.mhlw.go.jp/content/10800000/000749617.pdf (accessed 2021.3.01).

Japanese Ministry of Labour Health and Welfare (2020). About BCP guidelines for care facilities, Office communication from directors, 1214-1 (in Japanese).

Japanese Ministry of Labour Health and Welfare . (2019b). Survey results on the readiness of BCPs in hospitals (in Japanese) : https://www.mhlw.go.jp/content/10800000/000533729.pdf (accessed 2021.3.01).

Kanno, T., Ishida, C., Kanesaka, T., Okada, R., Kawazoe, T., Tae, K., & Sato, J. (2020). Resource-centric business continuity and resilience planning, *Proceedings of the 8th REA Symposium Embracing Resilience: Scaling up and Speed up*. *https://doi.org/10.15626/rea8.21*

ISO 22301:2019. (2019). International Standard, Societal security–Business continuity and management systems, Requirements.

A Systems-Resilience Approach to Technology Transition in High-Consequence Work Systems

Kelly NEVILLE, Heather ROSSO, and Bianica PIRES
The MITRE Corporation

ABSTRACT
The rate at which new information technology fails to successfully transition is disconcertingly high. We propose a systems-centered technology transition approach grounded in complex systems science and resilience engineering. In an exploratory effort, we developed an initial version of the Transform with Resilience during Upgrades to Socio-Technical Systems (TRUSTS) Framework. The framework, derived from a literature review and case-study assessments, specifies general characteristics of complex systems that give them the ability to perform resiliently in complex, high-stakes operations. The framework is being used to derive transition-support tools and metrics. Our goal is to use the TRUSTS Framework and tools to improve the success rate of technology transition where transition is counted as successful not just because the new technology is being used but because it transitioned without compromising the receiving system's ability to perform and adapt in high-demand conditions. Numerous high-consequence sectors will benefit, including transportation, healthcare, and national security.

KEYWORDS
Resilience framework; technology transition; system modernization.

INTRODUCTION

The successful transition of information technology (IT) into operational use has been a challenging endeavour for decades (see, for example, Deal & Hoffman, 2010; Neville, Linde, Hoffman, Elm, & Fowlkes, 2008; Patterson, Millitello, Su, & Sarkar, 2016). In this effort, we focus on improving transition of new IT into high-consequence work systems such as surgical and other healthcare delivery systems, military operations, transportation systems such as air traffic control operations, and petrochemical and other process control operations. New IT intended for high-consequence work systems such as these frequently faces push-back and produces fear, distrust, and anxiety. If it does make it into the target system, it is often disregarded and worked around (e.g., Koopman & Hoffman, 2003). Misalignments with established system priorities, practices, relationships, and dynamics can wreak havoc within the system and compromise the system's ability to adapt responsively to demands and challenges—i.e., to behave resiliently. The system's compromised resilience, in turn, threatens the safety and well-being of humans and the system's ability to fulfil its mission or purpose (Finkelstein & Dowell, 1996; Patterson et al., 2016; Sherwood, Neville, McLean, Walwanis, & Bolton, 2020; Trist & Bamforth, 1951). We hypothesize that deficiencies in current transition practices can be overcome by improving the methods, tools, and even the paradigm for IT transition and system modernization.

The dominant paradigm for technology transition is technology centric. Technology is developed to add new functions to an established system or to improve the way existing functions are performed. Accordingly, the focus of technology development efforts tends to be on achieving those functions and establishing their reliable performance. Integration into the recipient system tends to focus only on the ability of the incoming technology to interoperate and communicate with other technology in the system. Little to no attention is given to other critical aspects of how that technology will interact with system elements, processes, relationships, and capabilities, or the ways they all work together during system operations. Nor is attention typically given to implicit priorities and values built into the system. Consequently, new technology may be programmed in ways that either does not acknowledge a system's priorities and values, or that reflect a different set of priorities and values.

Once new technology is introduced, the host system must weather negative consequences of misalignments and disruptions. Some misalignments and disruptions will reveal themselves only when the system becomes stressed; for example, when required to operate under high demands or unusual conditions. Yet, when new technology is introduced into high-consequence work systems, stakes are high. In some cases, national security is at stake. In many cases, human lives and well-being are at stake. It is no wonder that examples of failed technology transition and system modernization abound (e.g., Dwiveldi, Wastell, & Henriksen, 2015; Opendoor, 2019).

New technologies often have great promise and offer critical capability advancements. Their potential is delayed and diminished, sometimes permanently, and sometimes their transition simply never succeeds. We need tools, methods, and a new paradigm for effectively guiding the entry of new IT into high-consequence complex systems.

We propose a systems-centered transition paradigm. This paradigm will help high-level decision makers and engineering teams attend to the ways a new technology interacts with system characteristics that enable system resilience. These system characteristics, which we call *system resilience sources*, give complex, high-stakes systems their ability to perform with resilience in the face of surprises, threats, perturbations, extreme conditions, and other events and conditions that take systems to and beyond the edge of their competency envelope.

Resilience has been defined as the ability to anticipate, adapt, and respond to changes in the operational environment (Woods, 2006a). In this paper we present initial work toward developing a framework that defines system resilience as a set of system resilience sources that underlie the system's ability to anticipate, adapt, and respond. After presenting the framework, we discuss plans and initial steps toward translating the framework into transition-support tools and metrics.

THE TRUSTS FRAMEWORK

The *Transform with Resilience during Upgrades to Socio-Technical Systems (TRUSTS) Framework* specifies system characteristics that enable system resilience. These characteristics, derived from a literature review, allow responsiveness to moment-to-moment variations in the environment, situationally adaptivity to surprises and extreme conditions, and evolution in response to emergent dynamics and changes in the system and environment. The TRUSTS framework also addresses factors that influence the form and nature of system resilience sources. Below, we present those influencing factors and then describe the current set of system-level resilience sources.

Figure 1 presents an early framework-development artifact that specifies (1) an early set of hypothesized system resilience sources and (2) categories of factors that influence those sources. Brief descriptions of the represented categories of influencing factors are presented below, followed by descriptions of the current set of hypothesized system resilience sources.

Factors that Influence System Resilience Sources

Representation of System-Level Macrocognition

The labels Flexecution Layer and Sensemaking Layer to the right of the six boxes in the upper right quadrant of Figure 1 refer to models of macrocognitive activity (Hoffman, Klein, & Woods, 2019). *Flexecution* involves real-time goal recalibration and replanning (Klein, 2007). *Sensemaking* refers to the processes involved in recognizing and making sense of the current situation, projecting into the future, and comparing projection- and experience-based expectations with incoming information (Klein, Moon, & Hoffman, 2006). These activities, which can be performed by an individual, team, or entire work-system, are required if a system's resilience sources are to be employed effectively.

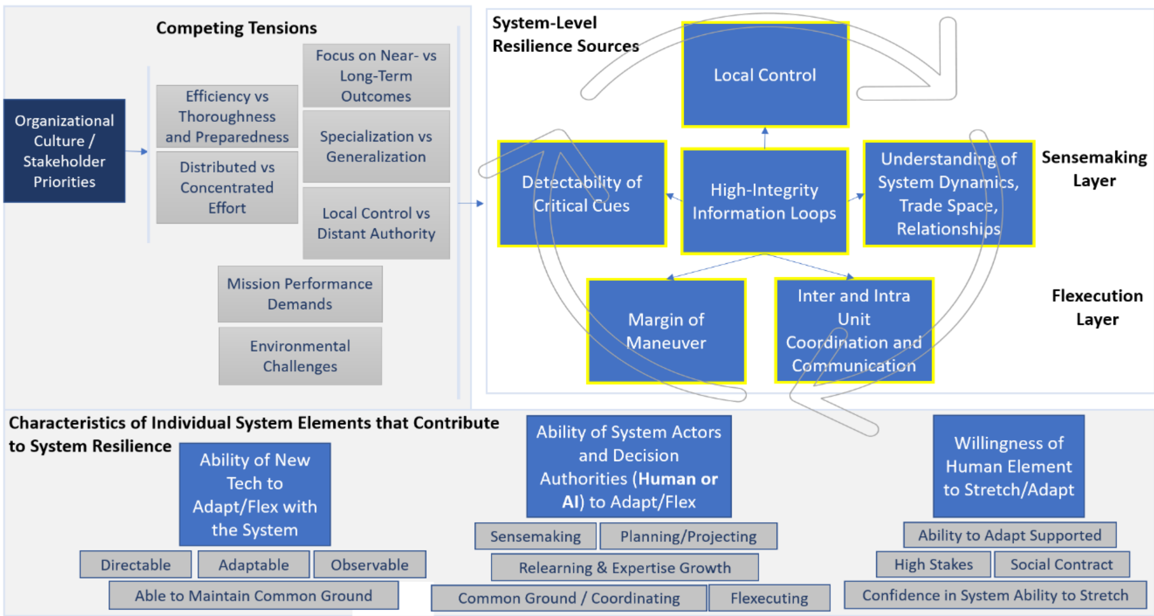


Figure 1. An early framework-development artifact. Categories of factors considered during the development of the TRUSTS Framework include system-level resilience sources (upper right corner; an early version of these sources is shown here), system-level macrocognitive processes such as sensemaking and flexecution (labels on the far right), competing tensions (upper left corner), and characteristics of individual system elements (bottom row).

Representation of Competing Tensions

In the upper left corner of Figure 1 are boxes that represent competing tensions explored in the work of Woods and his colleagues (e.g., Hoffman & Woods, 2011; Woods & Branlat, 2011) that translate into trade-offs reflective of the system's value system and priorities. Trade-offs among conflicting operational demands, stakeholder priorities, and environmental stressors can directly affect the extent and quality of a system's resilience sources. For example, trade-offs that favor generalization, distributed effort, and responsiveness to environmental challenges benefit system resilience more than trade-offs that favor specialization, concentrated effort, and responsive to stakeholder interests, respectively. When new technologies are introduced into a system, they have the potential to change the balance in existing trade-off equations, thereby compromising system resilience sources and putting system operations into conflict with the system's established value set and priorities.

Representation of Individual System Elements that Contribute to System Resilience

Along the bottom of Figure 1 are boxes representing individual system elements, including the to-be-transitioned technology (far left) and agents within the system (intelligent technologies or humans; middle). Grey boxes under 'Ability of Tech to Adapt/Flex with the System' refer to variables that affect an individual technology's ability to contribute to the resilience of a complex system (Klein, Woods, Bradshaw, Hoffman, & Feltovich, 2004; Woods, 2005). Under 'Ability of System Actors and Decision Authorities to Adapt/Flex' are grey boxes specifying macro-cognitive processes by which agents and agent teams within the system contribute to system resilience (Hoffman, Klein, & Woods, 2019). Boxes under 'Willingness of Human Element to Stretch/Adapt' specify factors hypothesized to influence the willingness of the human elements of a system to participate in system responses to anomalous, high-stress, and potentially dangerous conditions.

Hypothesized System Resilience Sources

We are evolving the TRUSTS Framework via continued assessment of literature and case studies. In Table 1, to provide better visibility to the framework's evolution, we list an earlier set of hypothesized system resilience sources and categories of influencing factors (Column 1) adjacent to those that comprise the baseline version of the framework (Column 2).

Table 1. Changes in TRUSTS Framework elements during baseline framework development.

Early Representation of Framework Elements	Current Representation of Framework Elements
High Integrity Information Loops	High Integrity Information Feeds
Local Control	Polycentric Control Across Layers of Narrowing Control Spaces
	Adaptivity to Information Update Rate and Detail
Detectability of Critical Cues	Sensitivity to Critical Cues Signifying Threat or Boundary
Understanding of System Dynamics, Trade Space, and Relationships	Unchanged
Margin of Maneuver	Unchanged
Inter and Intra Unit Coordination and Communication	Bi-Directional Coordination
Competing Tensions	Added to the resilience source *Understanding of System Dynamics, Trade Space, and Relationships* (see description below)
System-Level Macrocognition: Sensemaking and Flexecution	System Sensemaking and Coordinated Flexecution
Characteristics and Capabilities of Individual System Elements	Unchanged

Below, we present the current set of system resilience sources that we propose to be at the core of a sociotechnical system's resilience. The set consists of three categories of system resilience sources: The system's control scheme, sources that enable system sensemaking, and sources that enable coordinated flexecution. These resilience sources, shown in Figure 2, contribute to resilience across time scales and during operations that range from tactical, with tight control loops, to strategic, with looser control loops, to support both situational and evolutionary adaptation. In the paragraphs that follow, we provide an overview of each system resilience source. We begin with the system resilience source High Integrity Information Feeds, which crosses all three categories of resilience sources.

High Integrity Information Feeds is a system resilience source that benefits all other resilience sources. Feeds of environmental and system information are critical to the effectiveness of the other resilience sources. The extent of the benefit to other resilience sources is dependent on the integrity of the information. In particular, they are benefitted to the extent information is highly perceptible, clear, unobscured by visual or auditory clutter, accurate, reliable, coherent, understandable, timely, and with low latency.

The Control Scheme

Polycentric Control Across Layers of Narrowing Control, refers to a control scheme influenced particularly by Woods (Woods, 2006b; Woods & Branlat, 2011; Woods credits Ostrom [1990, 1999, 2003] with the polycentric control concept) designed to optimize a system's ability to exert local control, i.e., the ability to respond directly and effectively to variety and demands in the performance environment (Woods & Branlat, 2011). In this control scheme, units of a work system exert their influence in ways that cooperatively benefit local control. Two important forms of unit cooperation are as follows:

- Units at higher echelons monitor broadly and focus on system strategy. The results of their broad monitoring and strategic decision making are used to guide and reduce the breadth of the problem space for which organizationally lower units are responsible. This benefits the system's ability to exert local control. Strategic units enable greater local control by handling strategic sensemaking and decision making so that local-control units can focus on time-pressured tactical sensemaking and decision making.
- Units at all echelons overlap in their sensemaking and decision-making roles, thereby positioning themselves to quickly step in and assist or potentially take over when another unit experiences high-stress demands or falters.

In addition to the work system control structure, local control is dependent on:

- the quality and timing of information and feedback.
- access to flexible resources and response options.
- an appropriate balance between distant authority and local control. When variety in the operating environment is very high and rapid responses, improvisation, and resourcefulness are required—even for occasional unanticipated or novel conditions—strategic-level guidance must be loose (e.g., Flach et al., 2013). Guidance should inform but not restrict options as variety in control options must at least match the variety in internal or external conditions (Ashby, 1956).
- the ability of intelligent agents in the system to guide work-system adaptation and self-reorganization with resourcefulness and wisdom, which depends partly on training, knowledge, and experience.

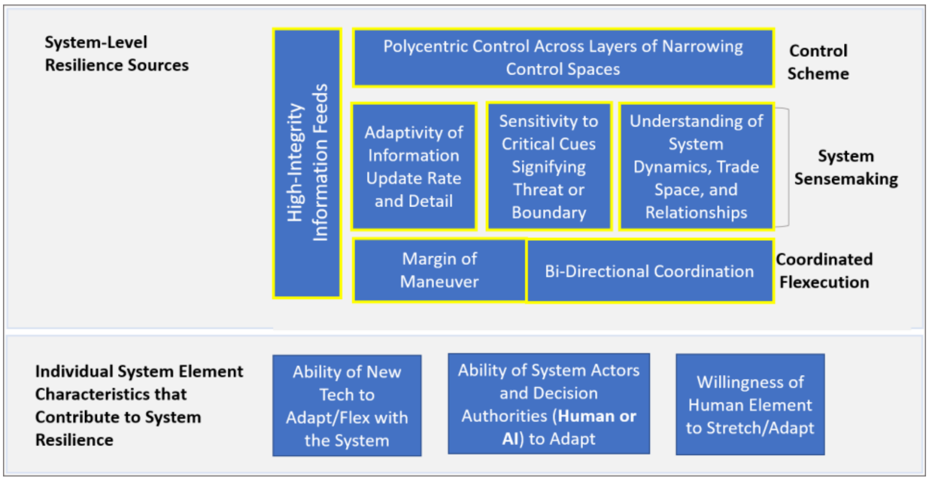


Figure 2. The TRUSTS Framework system resilience sources are shown in the upper half of the figure. Competing tensions have been incorporated into the system resilience source Understanding the System, and specifically, into understanding the system trade space, as described below.

System Sensemaking

Adaptivity of Information Update Rate and Detail ensures that work-system units at both strategic and tactical echelons receive information and feedback at rates and with levels of detail consistent with the speed and extent of change in the operating environment and system. When a work system transitions from routine to nonroutine operating conditions, the nature of information updates needs to transition as well. This system resilience source was suggested by a case study of a transportation firm's response to Superstorm Sandy (Deary, Walker, & Woods, 2013). In this study, strategic units of the work system did not have a means to track rapid change in the operating environment. Nor were information sources tuned to emergent system goals, such as protecting client assets. To cope, these units demanded frequent updates from operators at the local-control level, interfering with the ability of those operators, who were operating at the edge of their competency envelope, to maintain local control.

Understanding of System Dynamics, Trade Space, and Relationships refers to the need for system units to have system-level awareness and understanding in three key areas:

- *Interdependencies with Other System Units.* Each system unit should understand how its status and actions affect other units, what resources it can offer other units experiencing high-stress conditions, and what resources those units may be able to offer when it experiences high-stress conditions.
- *Decentralized, Cross-Function Communication of Deviations*. Each system unit should be given awareness of deviations from routine conditions in other units and in the system as a whole and deviations that may compromise system resilience. This includes deviations in operating conditions, resources available, protocols, and decision-making.
- *The System Trade Space*. Cross-unit awareness of the work system trade space enables a system to monitor its resilience and maintain coordination across units. Trade-space awareness encompasses:

- Measures or other forms of feedback representing competing pressures. Examples include measures of both efficiency and thoroughness and measures sensitive to both meeting stakeholder expectations and coping with challenges of the operational environment. These sources of feedback allow the work system to self-manage alignments between its performance and priorities and to detect misalignments between priorities and operational constraints. They also allow the system to detect unintentional shifts in performance relative to priorities.
- Deviations from routine trade-offs and priorities during nonroutine operations. Communicating these deviations across system units facilitates system-wide coordinated adaptation.

Sensitivity to Critical Cues Signifying a Threat or Boundary refers to mechanisms that enable a system to detect potential threats and detect them early before the threat has an opportunity to degrade system function. The system requires early cues about escalating performance conditions so that it can mobilize resources and anticipate, prepare, and begin to initiate courses of action. Early notification means the system has more time to adapt, and thus is critical to resilience. Sensitivity to early critical cues requires mechanisms for critically evaluating trends, feedback, and other incoming information. It requires a rich representation of 'normal' to facilitate the detection of unexpected non-normal events and means for recognizing known threats such as the approach of thresholds or conditions that require system adaptations. It requires a judicious use of automated alarms in favor of mechanisms that build awareness. It requires communication throughout the system. When critical cues are detected by any unit of a system, the system as a whole should be notified and should prepare to mobilize in response. Some units, for example, may prepare for direct contact with a threat; others may prepare to be in an on-call status.

Coordinated Flexecution

Bi-directional Coordination helps system units maintain common ground about system priorities, trade-off decisions, system-wide protocols, and more. The bi-directionality of coordination lines helps to ensure that the vertical influence exerted by strategic-level echelons is, via feedback, adjusted to align to the constraints and demands of the operational environment. Bi-directional resilience additionally aids in the flexibility of system responses to perturbations and challenges. A threat or anomaly may initially impact just one or two system units, or even just a portion of one system unit. The affected unit(s) use coordination lines to communicate adaptations and threat responses to other units. This paves the way for those other units to respond with resources, back up actions, preparations for possible ripple effects, and adaptations to compensate for the degraded part of the system. (This cross-unit cooperation is part of the notion of polycentric control.)

At the longer, strategic timescale, upward communication of unit-level changes to strategic units may allow those changes to be more effectively communicated and coordinated across the system. Strategic units can help the system learn and evolve by providing system-wide updates about priorities, resources, adaptations, and decisions.

Margin of Maneuver refers to "a cushion of potential actions and additional resources that allows the system to continue functioning despite unexpected demands" (Woods & Branlat, 2011, p. 281). In the TRUSTS Framework, we specify as contributing factors the variety and malleability of resources, roles, and response strategies. Stephens, Woods, Branlat, and Wears (2011) suggest another important factor is the ability to actively create and maintain adequate margins both within a given unit and by drawing on other units' margin, either cooperatively or defensively. Strategies for creating and maintaining margin they observed in a hospital system include (p. 257):

- defensive strategies that maintain local margin by restricting other units' actions or borrowing other units' margin(s),
- autonomous strategies to create margin via exclusively local reorganization, and
- cooperative strategies that either jointly recognize and manage existing common pool resources or create new common pool resources that allow sharing of margin with other units.

We noted a fourth strategy for margin creation in an ongoing case study of coordination among units of an air traffic flow management system: Initiation of a progressive-action feedback loop far in advance of the threat (a weather system in this case) interacting with the system. This strategy of initiating early action has the benefit of reducing the complexity of the approaching high-stress potential event. Each action produces early feedback that permits probing and adjustment to the course of action based on feedback from the operating environment. It allows the system to adjust its operations and dynamics in anticipation of high-stress conditions while they have the time to focus on internal adjustments. Thus, early and progressive action can prepare the system, adapt the course of action, and reduce the variety and complexity of the threat event. This type of preparatory loop is also mirrored in Snowden and Boone's (2007) concepts of *safe-to-fail experiments* and experimental probes followed by observation to determine the next step taken.

NEXT STEPS

Although we are in early states of developing the TRUSTS Framework, we consider its resilience sources to be sufficiently supported by research and theory to begin identifying and deriving transition-support tools and measures that are grounded in the framework. A brief overview of our plans and early progress follows.

Framework Development

Our next steps in framework development include (1) assessing whether to expand hypothesized system resilience sources to include social factors, (2) conducting an evaluation of the framework using case studies described in the research literature and interviews, and (3) using the case studies to characterize relationships among system resilience sources. The case-study-based evaluation of the TRUSTS resilience sources and their relationships is planned, and data collection is about to ensue. The work of assessing whether and how to incorporate social factors is underway. We provide an overview of the latter below.

One aspect of the social component of innovation adoption is willingness of humans in the system to stretch and adapt to incoming technology and innovation. Without the willingness of the human component, newly integrated technology can be sabotaged through workarounds and deliberate or unintentional misuse. Innovation also affects and is affected by the human component, affecting the resilience of the system as a whole. The TRUSTS Framework specifies willingness as a peripheral component of the TRUSTS Framework. We explore the role of the willingness component through the SCARF model of human social experience (Rock, 2015), which addresses assessment of social interrelations broken into five domains — Status, Certainty, Autonomy, Relatedness, and Fairness. As described by Rock (2015, p. 44)

- Status is about relative importance to others.
- Certainty concerns being able to predict the future.
- Autonomy provides a sense of control over events.
- Relatedness is a sense of safety with others, of friend rather than foe.
- Fairness is a perception of fair exchanges between people.

The resulting perceptions from these domains as defined by the SCARF model are “primary reward” or “primary threat”: “For example, a perceived threat to one’s status activates similar brain networks to a threat to one’s life. In the same way, a perceived increase in fairness activates the same reward circuitry as receiving a monetary reward” (Rock, 2015, p. 1). We set out to see if there was space for addressing the human, motivational components of systems by addressing SCARF within the TRUSTS Framework.

Using a case study of coal mining innovation (Trist & Bamforth, 1951) to explore possible connections between SCARF and TRUSTS factors, we identified shifts in SCARF factors before and after the introduction of innovation. One example of the effects of innovation on the human factor is the shift from small-group collaboration and coordination encompassing the “entire cycle of operations within the compass of its membership” (Trist & Bamforth, 1951, p. 6) to a factory-inspired assembly line structure. While an assembly line structure provided Certainty (a SCARF factor) in terms of a procedural structure and timeline, it resulted in diminished resilience in Coordination and Margin of Maneuver (TRUSTS factors), which subsequently resulted in diminished perceptions of Status (“Only in relation to this total cycle group could various smaller sub-groups secure function and acquire social form”; Trist & Bamforth, 1951, p. 10) and autonomy (“the spatio-temporal structure imposed…makes a difficult habitat when considered as a theatre in which effective communication and good working relationships must be maintained”; Trist & Bamforth, 1951, p. 14). It was discovered that system changes affecting TRUSTS’ Local Control and Margin of Maneuver factors also affected SCARF’s Autonomy, Certainty, and Status factors. The goal was to explore whether relationships existed between SCARF factors and work-system resilience represented by the factors of the TRUSTS Framework, and whether SCARF factors or adaptations should be added to the hypothesized TRUSTS Framework. This resulted in several observations:

- New-technology impacts on certain system-resilience sources tended to be correlate with impacts on certain SCARF factors. For example, effects on Status and Certainty co-occurred with effects on Coordination and Margin of Maneuver.
- We assessed the SCARF factors as relevant to system resilience. Specifically, we assessed them as contributing to the willingness of system operators to endure high workload and stress and help the work system survive and adapt to abnormal, edge, and extreme performance conditions.
- Poorly transitioned technologies may compromise the system-level resilience sources while simultaneously lessening the willingness of humans or intelligent agents to invest themselves in the process of adaptation to demanding conditions. SCARF factors will be added to the framework element Willingness of the Human Element to Stretch/Adapt.

Transition-Support Tools and Metrics

We are in the very early stages of deriving methods and tools to help decision makers and technology developers attend to how their technology designs and implementation strategies interact with system resilience sources. Our initial efforts are focused on identifying existing analysis tools and metrics for evaluating system resilience both holistically and by individual system resilience source. We are mapping tools and metrics to the TRUSTS Framework and its elements to identify gaps in coverage and to evaluate existing tools and metrics in term of their relevance to guiding technology transition. From there, we plan to adapt existing or derive new tools and measures for guiding and improving the technology transition process.

Unsurprisingly, sources of resilience that are most difficult to measure quantitatively seem to be associated with the fewest assessment tools and measures. Existing approaches to resilience evaluation and measurement include quantitative measures of network characteristics (e.g., Schmid, Korn, & Stanton, 2020), state- and phase-space analyses stemming from chaos and dynamical systems theory (e.g., Gorman, Amazeen, & Cooke, 2010; Holling, 1973), and holistic assessment frameworks such as the Resilience Analysis Grid (RAG; Hollnagel, 2011). Approaches from nonlinear dynamics show promise as they account for a system's multi-dimensionality as well as its inherent dynamics (Gorman et al., 2010). However, studies have not explicitly evaluated the sensitivity of such approaches to different sources of resilience. This limits their diagnosticity and leaves unknown their construct validity. By grounding tools in the TRUSTS Framework, we can achieve a comprehensive approach in which tools and metrics can be directly related to technology design and implementation decisions.

Other promising work includes efforts to develop qualitative system analysis and design tools that attend to system resilience sources. For example, Naikar (2020) has proposed the Work Organization Possibilities (WOP) modelling tool as an alternative to existing function allocation methods that disregard the benefits of role fluidity, resource sharing and co-opting, and flexibility in general to a system's margin of maneuver and ability to self-organize. Rankin, Lundbert, Woltjer, Rollenhagen, and Hollnagel (2014) developed a *strategies framework* for specifying strategies used to interpret and respond to variety in every day high-risk operations. It specifies strategies in terms of constraints, conditions, enablers, and action types. They developed a complementary *variety space diagram* for examining the extent to which the system's variety in sensemaking and control options enable a given strategy across performance environments featuring different amounts of variety. The variety space diagram thus offers a means to evaluate a system's margin of maneuver across operational contexts.

We are also exploring modelling-and-simulation (M&S) as a tool for exploring system resilience impacts of new technology. Because of its conduciveness to visualization, M&S provides a distinct advantage for communicating resilience concepts to decision makers and other stakeholders in ways that traditional methods and metrics fall short. Our hope is that simulations of system resilience sources will permit engineering teams to evaluate interactions between new technologies and the recipient system's resilience sources.

Our plans include developing federated simulation models instrumented with variables hypothesized to affect the state of the TRUSTS system resilience sources and sensitive to technology design and technology implementation characteristics. These models will allow decision makers and technologists to evaluate their technology designs and implementation plans. We also have plans to develop a simulation-based TRUSTS concept demonstration for communicating the framework to stakeholders.

CONCLUSION

Traditional engineering approaches tend to focus on maintaining system stability and robustness through preservation of capability and quick recovery to baseline performance (e.g., Madni & Jackson, 2009). In today's complex systems and operating environments, the system features that enable stability and robustness are important contributors to system resilience but inadequate on their own. The TRUSTS Framework emphasizes the critical roles that situational adaptivity and evolution play in the ability of complex systems to operate with resilience in complex and demanding environments. By specifying system characteristics that underlie a system's resilience, the TRUSTS Framework can help technologists and decision makers preserve and improve system resilience during technology transition and system modernization initiatives. TRUSTS-based tools and metrics can further guide and enable the conduct of system-centered technology transitions and upgrades.

We hypothesize that a system-centered approach to transition will be more productive than the typical focus on worker acceptance or management advocacy. TRUSTS aims to improve the success rate of technology transition and count transitions as a success not just because the technology is being used, but because it has transitioned without compromising the recipient system's ability to perform and adapt in high-demand conditions. We have begun investigating use of the TRUSTS framework to benefit airspace, healthcare, and military operations, and plan to use it to benefit supply-chain, process-control, network, and security system design and modernization.

ACKNOWLEDGMENTS

We are grateful to Robert Hoffman and David Woods feedback on the framework; Saurabh Mittal, Doug Flournoy, Kris Rosfjord, Matthijs Broer, and Beverly Wood for feedback on this paper, and Kris Rosfjord for her guidance and support. The views expressed in this paper are those of the authors and do not represent the official views of the organizations with which they are affiliated. An author's affiliation with The MITRE Corporation is provided for identification purposes only and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the author.

REFERENCES

Ashby, W.R. (1956). *An introduction to Cybernetics*. London: Methuen.

Binder, C. R., Mühlemeier, S., & Wyss, R. (2017). An indicator-based approach for analyzing the resilience of transitions for energy regions. Part I: Theoretical and conceptual considerations. *Energies, 10*(1), 36.

Deary, D. S., Walker, K. E., & Woods, D. D. (2013). Resilience in the face of a superstorm: a transportation firm confronts hurricane sandy. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 57, No. 1, pp. 329-333). Los Angeles, CA: Sage.

Deal, S. V., & Hoffman, R. R. (2010). The Practitioner's Cycles, Part 1: Actual World Problems. *IEEE Intelligent Systems, 25*(2), 4-9.

Dwivedi, Y. K., Wastell, D., & Henriksen, H. Z. (2015). Guest editorial: Grand successes and failures in IT: Private and public sectors. *Information Systems Frontiers, 17*(1), 11-14.

Flach, J. M., Steele-Johnson, D., Shalin, V. L., & Hamilton, G. C. (2013). Coordination and control in emergency response. In A. Badiru & L. Racz (Eds), *Handbook of Emergency Response: A Human Factors and Systems Engineering Approach* (pp. 533-548). Boca Raton, FL: Taylor & Francis.

Finkelstein, A., & Dowell, J. (1996, March). A comedy of errors: the London Ambulance Service case study. In *Proceedings of the 8th International Workshop on Software Specification and Design* (pp. 2-4). IEEE.

Gorman, J. C., Amazeen, P. G., & Cooke, N. J. (2010). Team coordination dynamics. *Nonlinear dynamics, psychology, and life sciences,* 14(3), 265.

Hoffman, R.R., Klein, G., & Woods, D.D. (2019). The theoretical foundation of macrocognition. Proceedings of the International Conference on Naturalistic Decision Making 2019, July 16-20, San Francisco, CA.

Hoffman, R.R, & Woods, D. D. (2011). Beyond Simon's slice: Five fundamental trade-offs that bound the performance of macrocognitive work systems. *IEEE Intelligent Systems 26*(6), 67-71.

Holling, C. S. (1973). Resilience and stability of ecological systems. Annual review of ecology and systematics, 4(1), 1-23.

Hollnagel, E. (2011). *RAG-The resilience analysis grid. Resilience engineering in practice. A guidebook.* Farnham, UK: Ashgate, 275-296.

Koopman, P., & Hoffman, R. R. (2003). Work-arounds, make-work, and kludges. *IEEE Intelligent Systems, 18*(6), 70-75.

Klein, G. (2007). Flexecution as a paradigm for replanning, part 1. *IEEE Intelligent Systems, 22*(5), 79-83.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent systems, 21*(5), 88-92.

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a" team player" in joint human-agent activity. *IEEE Intelligent Systems, 19*(6), 91-95.

Madni, A.M., & Jackson, S.E. (2009). Towards a conceptual framework for resilience engineering. *IEEE Systems Journal, 3*(2), 181-191.

Neville, K., Hoffman, R. R., Linde, C., Elm, W. C., & Fowlkes, J. (2008). The procurement woes revisited. *IEEE Intelligent Systems*, *23*(1), 72-75.

Opendoor (February 20, 2019). *The Standish Group report 83.9% of IT projects partially or completely fail.* Opendoorerp.com.

Patterson, E. S., Militello, L. G., Su, G., & Sarkar, U. (2016). Characterizing a naturalistic decision-making phenomenon: loss of system resilience associated with implementation of new technology. Journal of cognitive engineering and decision making, 10(3), 229-243.Rock, D. (2008). SCARF: A brain-based model for collaborating with and influencing others. *NeuroLeadership journal, 1*(1), 44-52.

Rankin, A., Lundberg, J., Woltjer, R., Rollenhagen, C., & Hollnagel, E. (2014). Resilience in everyday operations: a framework for analyzing adaptations in high-risk work. Journal of Cognitive Engineering and Decision Making, 8(1), 78-97.

Schmid, D., Korn, B., & Stanton, N. A. (2020). Evaluating the reduced flight deck crew concept using cognitive work analysis and social network analysis: comparing normal and data-link outage scenarios. *Cognition, Technology & Work*, *22*(1), 109-124.

Sherwood, S.M, Neville, K.J., McLean, A.L.M.T., Walwanis, M.M., & Bolton, A. (2020). Integrating new Technology into the complex system of air combat training. In H. Handley & A. Tolk (Eds.), A framework of Human System Engineering: Applications and case studies. Washington, D.C.: IEEE.

Snowden, D. J., & Boone, M. E. (2007). A leader's framework for decision making. *Harvard business review, 85*(11), 68-76.

Stephens, R. J., Woods, D. D., Branlat, M. & Wears, R. L. (2011b). Colliding Dilemmas: Interactions of Locally Adaptive Strategies in a Hospital Setting. *Proceedings of the Fourth Resilience Engineering Symposium* (pp. 256-262), June 8-10, 2011, Sophia-Antipolis, France.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human relations, 4*(1), 3-38.

Woods, D.D. (2005). Generic support requirements for cognitive work: laws that govern cognitive work in action. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 317-321). Los Angeles: Sage.

Woods, D. D. (2006a). Essential characteristics of resilience. In. E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 21-34). Ashgate.

Woods, D. D. (2006b). How to design a safety organization: Test case for resilience engineering. In. E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience engineering: Concepts and precepts* (pp. 315-325). Ashgate.

Woods, D. D., & Branlat, M. (2011). How human adaptive systems balance fundamental trade-offs: Implications for polycentric governance architectures. *Proceedings of the Fourth Resilience Engineering Symposium* (pp. 277-283), June 8-10, 2011, Sophia-Antipolis, France.

Designing for "Agility" in Envisioned Worlds: Concepts for Collaborative Intelligence in Human-Machine Teams

Neelam NAIKAR, Glennn MOY, Hing-Wah KWOK and Ashleigh BRADY
Defence Science and Technology Group, Australia

ABSTRACT
This paper explores two challenging design problems—that of designing "agile" sociotechnical systems, which can contend with very high levels of instability, uncertainty, and unpredictability, and that of designing future or envisioned systems, which do not yet exist. We suggest a design approach focusing on constraints or boundaries of successful operation, which can accommodate many possibilities for action, thereby providing a systematic basis for creating "agility" in the future system for dealing with dynamic, ambiguous environments. Given recent advances in artificial intelligence, we demonstrate some of these concepts with the decision ladder template from cognitive work analysis, showing how it may be applied in the design of "agile" human-machine intelligence. This design approach, which moves beyond standard approaches to function allocation, whether fixed or dynamic, toward 'team player', joint cognitive systems, and collaborative approaches to human-machine performance, has significant implications for research in artificial intelligence.

KEYWORDS
Adaptation; engineering; function allocation; human-machine teaming; cognitive work analysis; artificial intelligence.

INTRODUCTION

Recent events, such as the Australian and American wildfires and the global pandemic, have demonstrated how difficult it is to prepare appropriate actions, structures and tools for responding to complex unfolding and unexpected events.. People have been required to make decisions in the face of great uncertainty or incomplete information, for instance about how the COVID-19 virus spreads, and in spite of the inherent unpredictability of the evolving situation, such as where a fire might head next. How do people and organisations operate successfully in such situations, and how can we design, not just existing systems, but also future systems to withstand and exploit such dynamic, uncertain, and unpredictable conditions?

In this paper, we examine the problem of designing for "agility" in envisioned worlds, in the context of military command and control. We briefly describe the Australian Defence Force's future concept for command and control, which motivates this research, and explain what "agility" signifies in this context, noting its many similarities with concepts like adaptation and resilience. Subsequently, we consider some of the challenges of designing for agility in an envisioned world, and discuss an approach based on cognitive work analysis, which focuses designers' attention on the constraints or boundaries of successful operation (Rasmussen, Pejtersen & Goodstein, 1994; Vicente, 1999). We then demonstrate some of these ideas in relation to designing collaborative intelligence in human-machine teams, with the intent of enhancing agility in organisations. This focus recognises recent advances in artificial intelligence (AI), which greatly expand the possibilities for partnerships between humans and machines, but also introduce new challenges for agility in the wider sociotechnical system. New concepts are needed around the design of human structures, machine intelligence, and the interactions between these.

"AGILITY" IN FUTURE COMMAND AND CONTROL

The Australian Defence Force's (2018) vision for future command and control focuses on the period beyond 2035. This concept anticipates the need to alter Australia's methods for command and control because of projected changes in the future operating environment. A critical change that is envisaged is the need to shift from a hierarchical command and control structure to an organisational concept with greater levels of agility, particularly one with more fluid control structures for responding to complex unfolding events. This change is assessed to be needed in view of the nature of potential adversaries in the future environment, as well as a wide range of emerging technologies, some of which are expected to be widely available commercially.

The term *agility* in this vision for future command and control has many commonalities with the concepts of *adaptation* (Vicente, 1999) and *resilience* (Hollnagel, 2016) in the cognitive systems engineering and resilience engineering fields, respectively. Hollnagel has recently defined a resilient system as one that "can adjust its functioning prior to, during, or following events (changes, disturbances, and opportunities), and thereby sustain required operations under both expected and unexpected conditions.”. In the command and control field, *agility* has been defined as “the capability to successfully effect, cope with and/or exploit changes in circumstances” (NATO Research and Technology Organization, 2014, p. 54). Generally, both definitions allude to the capacity of a system to manage change successfully, and encompass the ability of the system to do so either before or after the change happens.

In this paper, we adopt a design perspective that focuses the discussion of agility on designing sociotechnical systems that are capable of contending with dynamic, uncertain, and unpredictable environments. This perspective not only recognises the reality of ongoing change or instability in the work environment, but also emphasises the likelihood of considerable uncertainty regarding many features or details of the changing circumstances. Further, it emphasises the inherent unpredictability of events or occurrences in complex environments, like military operations, which is widely recognised as posing the greatest threat to system performance and safety (e.g., Leveson, 1995; Vicente, 1999).

Given the inevitability of change, uncertainty, and unpredictability, we place the design focus on the concept of boundaries of successful operation (Figure 1). These boundaries are limits on behaviour in a system of work that must be respected by actors, regardless of the circumstances that are experienced or anticipated, for the system to perform effectively (Rasmussen et al., 1994; Vicente, 1999). Such boundaries may be delineated by physical or natural laws, such as laws of aerodynamics, or by human intentions, conventions, or values, such as principles of armed conflict. Thus, these boundaries place constraints on actors’ behaviours, whether human or machine, but within these constraints, actors still have many possibilities for action. Consequently, by developing designs that focus on such constraints, the intent is to allow the human-machine system to adjust its behaviours to accommodate local contingencies, even when unanticipated, and to accommodate workers’ preferences or local norms and cultural practices, without crossing the boundaries of successful operation. The value of this design approach is demonstrated by a large body of experimental studies (Vicente, 2002) and by industrial case studies (Elix & Naikar, 2020; Naikar, 2013).

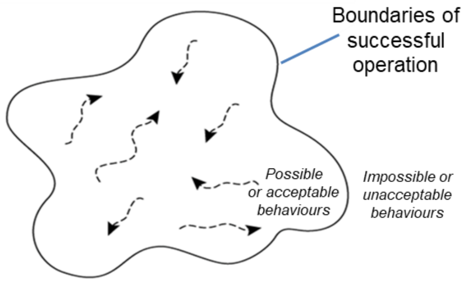


Figure 1. Within the boundaries, actors have many possibilities for behaviour (dashed arrows). From Naikar (2013).

ENVISIONED WORLD

The challenge of designing an agile command and control system, one that can contend with considerable instability, uncertainty, and unpredictability, is compounded by the fact that, in this case of the Australian Defence Force, the focus is on a future system. In other words, the object of design is a sociotechnical system that does not yet exist in the envisioned form, but that is thought to be needed for contending with the future task environment. The challenge therefore is one of designing for an envisioned world.

The envisioned world problem was articulated by Woods and Dekker (2000) in the context of their studies of future air traffic management systems. They identified four challenges confronted by analysts or designers (pp. 278-279). The challenge of *plurality* recognises that there are multiple perspectives of how proposed changes will affect the nature of work practice, held by different stakeholders. *Underspecification* recognises that each version of the envisioned world is necessarily a simplified or partial representation of “what it would mean to function in the field of practice in the future”. *Groundedness* recognises that the envisioned concept may “easily be disconnected from, and even contradict,” empirical research on people, technology, and work. Finally, *overconfidence* recognises that advocates of the envisioned concept are often overly confident that, “if the systems envisioned can be realised, the predicted consequences and only the predicted consequences will occur”.

Notably, this seminal articulation of the envisioned world problem (Woods & Dekker, 2000) is motivated by technological change, and questions about how advances in technology may dramatically transform a field of practice in the future. Where organisational change is mentioned, the implication is that such change is motivated by, or an inevitable result of, projected changes in technology. Certainly, given the pace of technological change, this is the most likely scenario that analysts and designers will confront and, in accordance with this observation, prior studies of envisioned worlds have been motivated by cases of prospective technological advances (e.g., Elix & Naikar, 2020; Militello et al., 2019).

In contrast, the agile command and control concept represents a somewhat different problem. That is, while this military concept is certainly shaped by anticipated technological developments, it is at least equally, if not predominantly, motivated by projected changes in military warfare given the characteristics of potential threats or adversaries, including their organisation. Consequently, changes to the command and control concept are deemed to be warranted that are fundamentally organisational in nature. This means that, aside from further studies of the organisational concept, a key challenge facing analysts and designers is the development of technologies that can enable agility in the future command and control organisation—including technologies that may not yet have been conceived or developed.

Another point of departure from the original articulation of the envisioned world problem is the methods adopted for addressing it. Woods and Dekker (2000) focus on descriptive methods of analysis and design, where specific future incidents are selected for study and behaviours and work practices in the envisioned world explored in this context. Such studies are definitely useful and needed in the study of envisioned worlds, but it is unclear to what extent they can serve as a basis for generating novel designs, rather than as a basis for evaluating alternative design concepts generated through other means. Another problem is that predicting situations of use is difficult, as Woods and Dekker acknowledge.

Further, in the envisioned command and control system, the technological possibilities are still to be defined, so that behaviours and work practices that may emerge in the envisioned world with people's continuing use of technologies, and increasing familiarity with the organisational concept, over time cannot be easily comprehended at this stage. Needless to say, waiting until the technologies are better defined is undesirable, as making changes to those technologies at later stages may be costly, perhaps prohibitively so. Moreover, it is well known that technologies intended to support people's current or observed behaviours generally lead to changes in those behaviours once those technologies are introduced into the work setting in a never-ending task-artifact cycle (Carroll and Rosson, 1992). For these and other reasons, therefore, alternative methods for designing for the envisioned world must also be considered.

In this paper, we focus on an approach to the envisioned world problem, which may be described as formative rather than descriptive in orientation (Rasmussen et al., 1994; Vicente, 1999). This approach entails designing for the envisioned world with the boundaries of successful operation as a key point of reference. As already discussed, these boundaries encompass many possibilities for behaviour, including those that may be necessary in unforeseen circumstances, which by definition cannot be observed or studied ahead of time. Such a design approach allows the human-machine system to evolve over time and to adjust or change its behaviours in response to situational demands, within certain boundaries. Known as cognitive work analysis (CWA), this approach has previously been applied to envisioned world problems with tangible impact on practice in several industrial cases (Elix & Naikar, 2020; Naikar, 2013).

In what follows, we focus on the goal of designing human-AI systems that are conducive to the aspirations of agility in an envisioned world, as articulated above. Recent advances in AI greatly expand the possibilities for partnerships between humans and machines, but also introduce new challenges for agility in the sociotechnical system. For example, new technologies, such as deep machine learning, pose significant opportunities for faster, more accurate and well-informed decision making. However, they are also comparatively brittle, often imposing pre-defined constructs or structures on a situation that may be inappropriate, particularly in unexpected scenarios. In practice, sociotechnical systems rely significantly on the inherent flexibility of humans and their ability to mould structures and technology when presented with different or unforeseen situations. By contrast, state of the art AI techniques are designed to act within known and pre-defined contexts, and often don't provide mechanisms for understanding or sharing human and machine intentions or accommodating fluid, ill-defined situations and their accompanying changes in human goals, assumptions, and priorities. Therefore, adoption of current AI systems will introduce many new challenges for human decision makers. If the sociotechnical system is not designed with agility as a priority, failure when under stress is likely. We therefore anticipate that agility will be a critical design factor in the command and control system of the future. To that end, we briefly consider some of the reasons that standard function allocation approaches underpinning studies of human-automation interaction may be unsuitable for designing for agility, before turning to the value of CWA for this purpose.

HUMAN-AUTOMATION INTERACTION

The literature on human-automation interaction is dominated by discussions of function allocation. Much of the conversation focuses on the levels of automation taxonomy (Sheridan & Verplank, 1978), and its subsequent variations incorporating stages of human information-processing (e.g., Parasuraman & Wickens, 2008), which characterise the distribution of function between human and machine in a number of levels ranging from fully manual to fully automated. Although the intent of research on this taxonomy has been to provide designers with guidance for automation design, by characterising patterns of situation awareness, workload, and performance associated with the different levels, a recent analysis has demonstrated that the taxonomy has little predictive value in complex work environments (Jamieson & Skraaning, 2018).

Outside the human factors field, many engineers and technologists appear not to be aware of the predictive intent of the taxonomy, seeing it simply as a framework for characterising different technological solutions. However, even for that purpose, the taxonomy may not be useful. In a report by the Federal Aviation Administration (Abbott, McKenney, & Railsback, 2013), difficulty was expressed in applying the taxonomy to characterise flight deck automation, to the extent that the taxonomy was abandoned. The model has also been rejected as a basis for designing human-automation interaction by a US Defense Science Board (2012). Further, while the Society of Automotive Engineers has recently adopted the taxonomy, their representation of the dynamic driving task appears to be a simplified abstraction that excludes active safety and driver assistance systems (Jamieson & Skraaning, 2018). Based on such examples, the practical experience seems to be that the taxonomy does not provide a faithful representation of many human-machine systems, inevitably leading to simplified representations that ignore the complexity of the design and problem.

A significant shortcoming of the taxonomy is that it is based on a flawed conceptualisation of how complex work is carried out in sociotechnical systems (Naikar, 2018). In particular, the taxonomy focuses designers' attention on individual operators, and on prescribing 'who should do what'—the human or the machine. However, in complex systems, people rarely, if ever, work in isolation. They operate in groups or teams, even if they are distributed in time or space, and many field studies show that they spontaneously adjust both their individual behaviours and their collective structures in a self-organising manner that moves the human-machine system toward becoming fitted or adapted to the demands of the evolving work environment (Naikar, 2020; Naikar & Elix, 2019). Attempts therefore to prescribe 'who should do what' are likely to lead to brittle designs that do not perform well in continually changing work environments, especially during events that have not been—and cannot be—anticipated.

Finally, another limitation of the taxonomy is that designers are not encouraged to think about the capabilities of humans in context of the work in situ, consequently promoting a 'left-over' approach to function allocation (Roth, Sushereba, Militello, Diiulio, & Ernst, 2019). Particular functions or tasks are automated because they can be, not necessarily because they will result in better performance. While it may seem logical to believe that if a task previously performed by humans can be performed by machines, gains in productivity, safety, or efficiency will be realised, the transformative effects of the technology on work practice are often not taken into account. Therefore, even if gains are observed in the targeted area, there may be unintended, undesirable effects elsewhere, so that overall system performance may not be improved and may in fact be compromised. For example, improved machine decision-support may lead to degraded awareness in humans or less flexibility in information structures.

COLLABORATIVE INTELLIGENCE

Given the problems articulated above with traditional approaches for human-automation interaction, we present some ideas for designing collaborative human-AI teams, with the intent of providing a formative approach for design based on the boundaries of successful operation. Our research on collaborative intelligence is motivated by the goal of creating systems with high levels of agility for future command and control, focusing on the period beyond 2035. Our ideas combine thinking about designing for self-organisation (Naikar, 2018, 2020; Naikar & Elix, 2019) with 'team player' (Christoffersen & Woods, 2002), joint cognitive systems (Woods & Hollnagel, 2006), and collaborative approaches (Roth, DePass, Harter, Scott, & Wampler, 2018; Roth et al., 2019) to human-automation interaction.

Designing Collaborative Human-AI Systems

In designing human-AI teams, we are motivated by many rich and comprehensive field studies that show that workers in sociotechnical systems adapt not only their individual behaviours but also their structures, or organisation, in line with the evolving situation (for a review, see Naikar & Brady, 2019). For example, emergency management workers constantly adjust their behaviours, improvising with tools, adjusting their standard routines, and even breaching standard operating procedures if necessary (Bigley & Roberts, 2001). Similarly, Rochlin, LaPorte, and Roberts (1987) found that the work organisation on naval aircraft carriers shifts spontaneously between formal and informal structures, in ways that are not planned a priori, centrally coordinated, or controlled through external intervention, such that the mapping between people and roles changes with the circumstances.

In this context, the concept of self-organisation (e.g., Haken, 1988) is important. Analysed in relation to sociotechnical systems (Naikar, 2020), this concept suggests that a system's structure may constrain its response in ways that are unsuitable or ineffective when particular circumstances are encountered. However, in responding to local conditions, individual, interacting actors may engage in spontaneous behaviours from which novel structures emerge that enable the system to respond appropriately to the circumstances. A new structure may be suitable for a time, constraining and enabling behaviours in ways that are appropriate to the circumstances—until the situation changes—and the spontaneous actions of individual, interacting actors results in further structural changes.

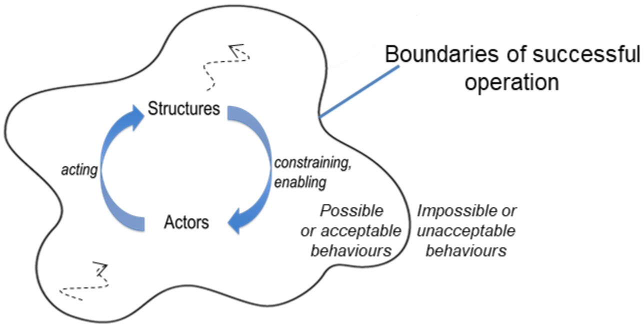


Figure 2. Illustration of the phenomenon of self-organisation in sociotechnical systems. From Naikar (2020).

Further, under conditions of instability, uncertainty, and unpredictability, it is proposed that workers move away from formal structures or procedures toward intrinsic constraints of the sociotechnical system as the principal governing mechanism for their conduct (Naikar, 2020). These constraints, or boundaries on behaviour, which must be respected by actors for a system to perform effectively, still afford actors many degrees of freedom for action. Therefore, given a commitment to these fundamental constraints on behaviour, rather than to formal specifications, actors can safely and productively adjust their actions to the local conditions, such that new structures may emerge from their spontaneous behaviours. This phenomenon of self-organisation is integral to managing instability, uncertainty, and unpredictability in the work environment.

In designing for self-organisation, the CWA framework provides a useful foundation (Naikar & Elix, 2019). This framework identifies the constraints or boundaries on successful operation in five different dimensions, specifically the constraints of the work domain, activity, strategies, social organisation and cooperation, and competencies. In the following section, we show how the decision ladder template from the activity dimension provides useful concepts for designing collaborative human-AI systems with high levels of agility. While other CWA dimensions are also important for our goal, the decision ladder template provides a more direct point of comparison for standard approaches to function allocation.

Prior to that, we acknowledge that the potential value of CWA in designing human-automation interaction has been recognised in a number of earlier studies. For example, Li and Burns (2017) suggest an approach that combines CWA with a formulation of the levels of automation taxonomy that incorporates stages of human information-processing. Nevertheless, as the focus of this approach is still on specifying ‘who should do what’, it is not wholly compatible with our intent of designing for agility. In another approach, Roth et al. (2019) focus on using CWA as a means for understanding and visualising the complexities of the work in context. Our approach builds on this perspective by more tightly integrating CWA with design, thereby recognising the value of the framework, not just in understanding the work in context, but also more directly in design.

Decision Ladder

Before discussing the value of the decision ladder for designing collaborative intelligence, it is important to emphasise that, although this tool is informed by observations of human behaviour in complex settings, the decision ladder is not a model of human decision making. Rather it is a template for design that accommodates both human and machine actors (Rasmussen et al., 1994; Vicente, 1999). Specifically, we view the decision ladder as a template that focuses attention on particular boundaries or constraints on successful operation that must be respected in the design of systems, whether by humans or machines. A generic depiction of the decision ladder is shown in Figure 3. The shunts and leaps in the centre of the figure are only *examples*, and are intended to emphasise that decision making in complex environments does not progress in a linear fashion from left to right, but is messy, spontaneous, and highly dependent on the situation. Thus, in designing collaborative intelligence, we see the decision ladder as providing a functional representation for design, much like a geographical map provides a functional representation for spatial navigation.

In contrast to standard approaches to function allocation, which focus design questions on ‘who should do what’, the decision ladder focuses attention on ‘who can do what’, framed in relation to boundaries of successful operation. This approach emphasises that both humans and machines may have potential for effective action within the boundaries in relation to the same work or activities (albeit perhaps using different strategies or processes). Consequently, ‘who does what’ (and how) at any point in time may vary spontaneously depending on the situation. For example, whether work is distributed or shared across human actors, machine actors, or both may depend on the nature of events, which may be unanticipated; the preferences or competencies of individual workers; or the speed and capacity of machine processing. Thus, the aim of design must be to maximise the organisational possibilities to create greater agility in the system.

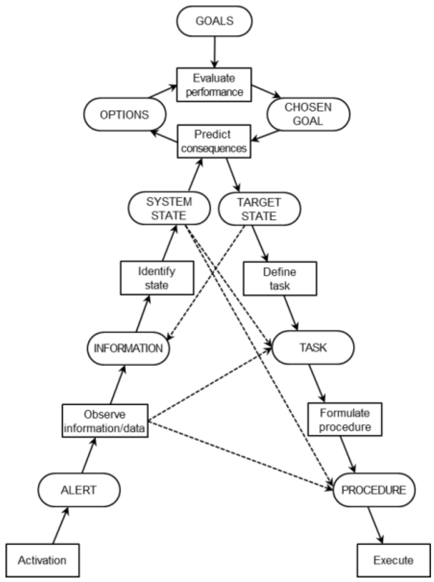


Figure 3. The decision ladder template.

Taking this design approach, the actual distribution of work between humans and artificial agents is emergent, given the constraints on their respective possibilities for action (Figure 4), some of which may be deliberately specified by designers, for example, for ethical reasons (such boundaries are the focus of social organisation and cooperation analysis in the CWA framework; see Naikar & Elix, 2016). Notably, this approach differs from both fixed and dynamic function allocation (e.g., Parasuraman & Wickens, 2008) in that it is not necessarily concerned with specifying a priori optimal allocation of functions between actors, whether humans or machines, given assumptions about how work should be distributed in relation to anticipated variations in situational parameters. Rather, it is concerned with bounding the possibilities for work organisation between human and artificial agents, with the actual distribution emerging in situ from their respective spaces of possibilities for action, which may be overlapping, thus creating redundancy in the system. This approach, which recognises that many events or contingencies and the details or nuances of those situations cannot be anticipated ahead of time by designers, is consistent with the intent of designing for agility.

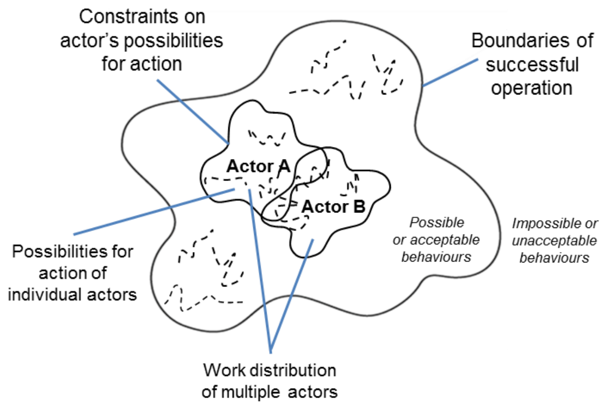


Figure 4. Given their possibilities for action, the work distribution between human and machine actors emerges in situ.

Table 1 extrapolates key boundaries from the decision ladder in terms that are relevant to command and control, specifically using the following labels: alert, information, situation, intent, options, plans, and actions. These labels are intended to capture constraints or boundaries on problem-solving, specifically in relation to the need to be attuned to signals, the need for information, the need to appreciate the evolving situation or context, the need to appreciate the intent, the need for options, the need to plan tasks, and the need to act. The table also shows some key design questions for designing collaborative human-AI systems, which recognise the potential of both humans and machines for effective action in relation to these boundaries. The intent of design then is not to allocate functions to either human or machine actors, but rather to support their respective spaces of possibilities for action, so that work can be shared or shift fluidly between the actors. These design questions should be considered in context of specific fields of work practice.

Table 1. Examples of design questions for collaborative human-AI teams.

Boundaries of Successful Operation	Questions for Identifying Bounds on the Human and Machine Actors in a System	
	Human Actors	**Machine Actors**
Alert: need to be attuned to signals	What can the human actors notice or detect?	What can the machine actors notice or detect?
Information: need for information	What can the human actors obtain or access?	What can the machine actors obtain or access?
Situation: need to appreciate the evolving situation and context	What can the human actors know or identify?	What can the machine actors know or identify?
Intent: need to appreciate the intent	What can the human actors infer or deduce?	What can the machine actors infer or deduce?
Options: need for options	What opportunities can the human actors see or find?	What opportunities can the machine actors see or find?
Plans: need to plan tasks	What can the human actors anticipate or predict?	What can the machine actors anticipate or predict?
Actions: need to act	What can the human actors do?	What can the machine actors do?

The decision ladder questions were inspired by, and complement, the work of Roth et al. (2018), who posed the following three questions in relation to the design of human-automation collaboration: (1) "What are the 'people on the scene' likely to 'know' that the automated aid cannot know?"; (2) "What situations are likely to arise that are beyond the capabilities of the automation working 'on its own'?"; and (3) "How best to draw on the knowledge and capabilities of the human and the automation to generate better 'joint' performance than would be possible with either one working alone?". The decision ladder framework builds on this approach, particularly the first question, by recognising explicitly the capabilities of both humans and machines, which Roth et al.'s questions allude to implicitly, and by expanding the set of design questions. Further, the decision ladder framework separates the capabilities of human and machines (who can do what) from what actually happens on the scene (who does what), depending on the nuances of the situation, including the preferences of individual workers, consistent with the goal of designing for agility. Lastly, in designing collaborative intelligence, we envisage utilising Roth et al.'s more detailed design questions about the kinds of visualisations and control mechanisms that may support collaboration between human and machine actors as a starting point.

IMPLICATIONS FOR DEVELOPING ARTIFICIAL INTELLIGENCE

We conclude this paper by considering the implications of the concepts presented above for future research on artificial intelligence. Technology solutions are typically developed with a well-defined customer need in mind. Whether through an upfront product-requirements capture, or through iterative development focused on customer needs, software development seldom emphasises agility as a core design requirement. In complex military contexts, users often underestimate the degree of flexibility in their own processes (e.g., Moy & Yesberg, 2014). In practice, this leads to technology being developed that significantly constrains the processes and options available to humans. In some situations, where processes are well-defined or routine, software can be optimised for these. However, in complex military situations such as planning and execution of military operations, there is often significant variability between events. This includes in the processes, information requirements and decision needs of the humans involved. In such contexts, inflexible machine processes often lead to low adoption rates for technology. This is underscored by reviews of intelligent systems developed for the military domain, which show that although many prototypes are developed, they are rarely fielded (e.g., Coombs, 2019).

We assess that artificial intelligence that supports military command and control systems need to be designed with process flexibility as a first-class design principle. In practice, this means that software development approaches that focus on a customer's own articulation of "how things are done" may be flawed, unless new emphasis is placed on the potential for novel or unexpected possibilities, within the broader boundary of successful operation. In addition, as the analysis above outlines, for agility in the wider sociotechnical system to be achieved, more research is needed to explicitly develop artificial intelligences with awareness of their capabilities *in the current context*. This introduces new research requirements for artificial intelligences to better articulate what they do and do not know in context and their ability to share this information with humans. While there is some work in this space (e.g., assessment of confidence of predictions, techniques for explainability), many emerging artificial intelligence technologies (e.g., deep neural networks) are often treated as a black-box, with limited ability to share information about the reasoning or limitations around their decisions. Significant research is required to enable many modern artificial intelligence technologies to efficiently share information related to the design questions presented above.

ACKNOWLEDGMENTS

Neelam Naikar thanks Dr Emilie Roth, Roth Cognitive Engineering, for several discussions about her ideas on designing for human-automation collaboration.

REFERENCES

Abbott, K., McKenney, D., & Railsback, P. (2013). *Operational use of flight path management systems*. Final Report, Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team.

Australian Defence Force. (2018). *ADF concept for command and control of the future force*. Commonwealth of Australia.

Bigley, G. A, & Roberts, K. H. (2001) The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, **44**(6), 1281-1299.

Carroll, J. M., & Rosson, M. B. (1992). Getting around the task-artifact cycle: How to make claims and design by scenario. *ACM Transactions on Information Systems*, **10**, 181-212.

Christoffersen, K., & Woods, D. D. (2002). How to make automated systems team players. In *Advances in human performance and cognitive engineering research* (pp. 1-12). Emerald Group Publishing Limited.

Coombs, G. (2019). *A review of intelligent decision support systems for land tactical planning*. DSTG Technical Note (DST-Group-TN-1921). Defence Science & Technology Group.

Defense Science Board. (2012). *The role of autonomy in DOD systems*. Office of the Under Secretary of Defense for Acquisition, Technology and Logistics.

Elix, B., & Naikar, N. (2020). Designing for adaptation in workers' individual behaviors and collective structures with cognitive work analysis: Case study of the diagram of work organization possibilities. *Human Factors*, **63**(2), 274-295.

Haken, H. (1988). *Information and self-organization: A macroscopic approach to complex systems*. Springer.

Hollnagel, E. (2016). *Resilience engineering*. https://erikhollnagel.com/ideas/resilience-engineering.html. Retrieved on 17 February 2021.

Jamieson, G. A., & Skraaning, G. (2018). Levels of automation in human factors models for automation design: Why we might consider throwing the baby out with the bathwater. *Journal of Cognitive Engineering and Decision Making*, **12**(1), 42-49.

Leveson, N. G. (1995). *Safeware: System safety and computers*. Addison-Wesley.

Li, Y., & Burns, C. M. (2017). Modeling automation with cognitive work analysis to support human-automation coordination. *Journal of Cognitive Engineering and Decision Making*, **11**, 299–322.

Moy, G., & Yesberg, J. (2014). *Lessons in software development for Defence from the JMAP planning tool*. DSTO Technical Report (DSTO-TR-2873). Defence Science & Technology Organisation.

Militello, L. G., Roth, E. M., Scheff, S., Ernst, K., Sushereba, C., & Klein, D. (2019). *Crew configuration analysis for future airborne reconnaissance operations*. Applied Decision Science.

Naikar, N. (2013). *Work domain analysis: Concepts, guidelines, and cases*. CRC Press.

Naikar, N. (2018). Human-automation interaction in self-organizing sociotechnical systems. *Journal of Cognitive Engineering and Decision Making*, **12**(1), 62-66.

Naikar, N. (2020). Distributed cognition in self-organizing teams. In M. McNeese, E. Salas, & M. R. Endsley, *Handbook of distributed team cognition, Contemporary research: Models, methodologies and measures in distributed team cognition*. Taylor & Francis.

Naikar, N., & Brady, A. (2019). Cognitive systems engineering: Expertise in sociotechnical systems. In P. Ward, J. M. Schraagen, J. Gore, & E. Roth (Eds.), *The Oxford handbook of expertise: Research & application*. Oxford University Press.

Naikar, N., & Elix, B. (2019). Designing for self-organisation in sociotechnical systems: Resilience engineering, cognitive work analysis, and the diagram of work organisation possibilities. *Cognition, Technology & Work*.

NATO Research and Technology Organization. (2014). *C2 agility* (SAS-085). Author.

Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors*, **50**(3), 511–520.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. John Wiley & Sons.

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), 76-90.

Roth, E., DePass, B., Harter, J., Scott, R., & Wampler, J. (2018). Beyond levels of automation: Developing more detailed guidance for human automation interaction design. *Proceedings of the Human Factors and Ergonomics Society 2018 Annual Meeting*, 150-154. Human Factors and Ergonomics Society.

Roth, E., Sushereba, C., Militello, L. G., Diiulio, J., & Ernst, K. (2019). Function allocation considerations in the era of human-autonomy teaming. *Journal of Cognitive Engineering and Decision Making*, **13**(4), 199-220.

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. Technical Report. MIT Man-Machine Laboratory.

Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Lawrence Erlbaum Associates.

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors*, **44**(1), 62-78.

Woods, D., & Dekker, S. (2000). Anticipating the effects of technological change: A new era of dynamics for human factors. *Theoretical Issues in Ergonomics Science*, **1**(3), 272-282.

Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. CRC Press.

Conflict Between Frames When Using Machine Learning

Hebah BUBAKR[a] and Christopher BABER[b]
[a] *The Department of Electronic, Electrical and Systems Engineering, University of Birmingham, HXB815@student.dham.ac.uk, on leave from King Faisal University, Saudi Arabia, Habubakr@kfu.edu.sa*
[b] *University of Birmingham*

ABSTRACT

Artificial intelligence and Machine learning (AI/ML) systems promise to improve organisational decision making by avoiding bias because the machine ought to remain unaffected by moods, prejudices or personal opinions when interpreting the data. However, this promise rests on the fact that these tools are independent of the biases of their developers. The purpose of this paper is to investigate what bias means to developers of AI / ML systems, and how they interpret bias through the result of the system. Our concern is with the relationship between developer - algorithm - data - output. In this paper, we applied the Data/Frame Model (DFM) to understand what decisions are made by developers of AI/ML. We propose that developers work with three distinct frames. First, they need to define a suitable dataset that will answer specific questions and also be amenable to analysis. We term this the 'dataset frame' and it includes factors such as size, representativeness and coverage, type of questions that could be addressed using these data. Second, having selected datasets, participants then explored different algorithms to test the selected datasets. We terms this the 'Algorithm Frame'. Third, once the algorithm produces answers, then these are reviewed. We terms this the 'Interpretation Frame' which includes both judgement on the performance of the algorithm (so overlaps with the 'algorithm frame'), plus judgement on the interpretation of the output to the original questions, and also judgement of the implications of this interpretation. Our conclusions suggest that developers of AI / ML might take a narrow perspective on 'bias' (as a statistical problem rather than a social or ethical problem). This is not because they are unaware of wider, ethical concerns but because the requirements relating to the management of data and the implementation of algorithms might narrow their focus to technical challenges. Consequently, biased outcomes can be produced unconsciously because developers are simply not attending to these broader concerns. This suggests that the 'interpretation frame' ought to be elaborated to encompass the implications arising from possible interpretations of the algorithms' output.

KEYWORDS

Judgment and Decision Making, Bias, Data frame model, Artificial intelligence, Machine learning

INTRODUCTION

Recent studies have focussed on explainable AI (Borgo et al., 2018; Anjomshoae et al., 2019; Baber et al., 2020, 2021) in which people interpret the output of the AI. The focus of these studies has been on developing our understanding of the ways in which people make use of the output of AI / ML systems. There has, we believe, been less attention given to the ways in which the systems are developed. Thus, our concern in this paper is with the relationship between developer - algorithm - data – output during the development of AI/ML. In order to explore this, we applied the Data/Frame Model (DFM) to understand what decisions are made by develolpers of AI/ML.

SENSE-MAKING AND DATA-FRAME MODEL

For Klein et al. (2006 a) "Sense-making is a motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively" [p. 71]. An important point to emphasise is that frames change as we add, change, delete or modify data. As shown in Figure 1, a frame allows a person not only to decide if data are sufficient, but to also project into the future what data will be needed. In this case, data can provide anchors on which to construct a frame which can then lead to search for further data. Klein et al. (2007) pointed out that the environment, individual characteristics, and available information will additionally influence the frame that is chosen. Once a frame has been selected, we seek data that relate to the frame. Any data that does not fit the frame is likely to be ignored or redefined. If the data are not suitable, then we might switch to a different frame.

Each element in the sense-making process forms relationships with others, using anchors, cues and any relevant data found in that environment. Thus, "when people try to make sense of events, they begin with some perspective, viewpoint, or framework" (Klein et al (2006 b). In terms of this perspective, there may be situations in which the output produced by AI might not be acceptable socially, at least, to the people who are affected by that decision. This might be due to failure of the algorithm, but more likely, it arises from the selection of data that impose the 'frame' upon which the algorithm operates. This means that the datasets might simply reflect biases and injustices that are prevalent in the society that produced these data (O'Neill, 2016). The issue that we explore in this paper is how do developers of AI / ML deal with the data and what sort of frames do they apply in their development and application of algorithms?

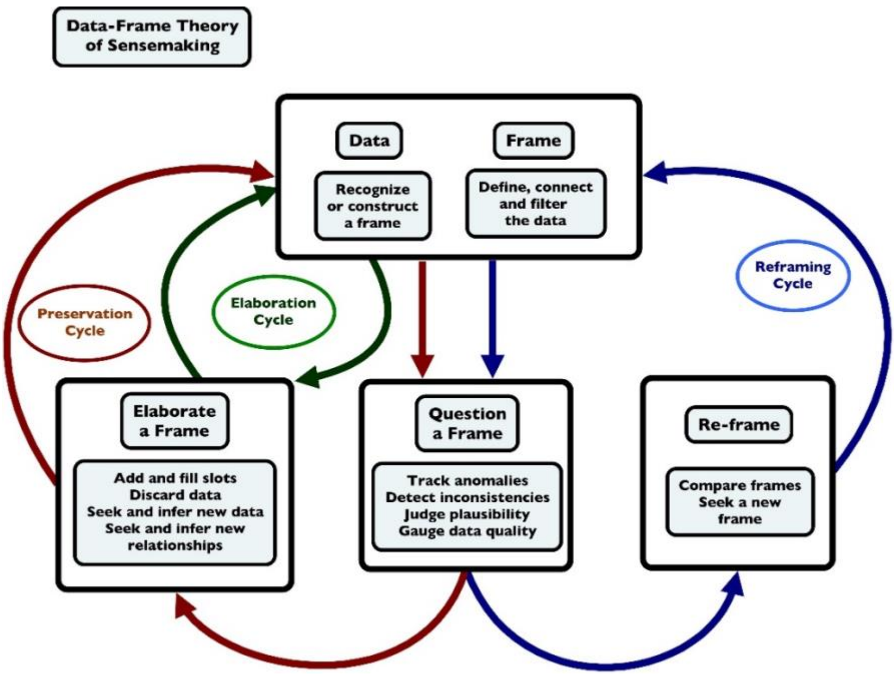


Figure 1: Data-Frame Model (after Klein et al., 2007)

METHOD

To understand how AI / ML developers work with their 'frames' we need to understand the process of creating an AI system. Interviews were conducted with Computer Science undergraduates who completed a project implementing Machine Learning on medical datasets. In order to encourage them to focus on the potential issues relating to 'bias' we devised a project which we called 'the bias machine'. The goal of the project was to deliberately produce biased output from datasets in a which could allow 'bias' to be dialled up or down, i.e., from no bias to high-levels of bias which would disadvantage particular groups of people. Using open access databases, participants selected two datasets to 1) make predictions about future smoking habits of different user groups based on racial profiles, and 2) classify severe visual impairment based on age, gender and race. From these datasets, participants would adjust datasets to inflict bias and implement the algorithms to produce biased results. After the AI / ML development was complete, participants gave a presentation on their work. Following this, we conducted Critical Decision Method (CDM) interviews with the participants, focusing on three high-level questions:

- What frame do developers apply when they select a dataset?
- What frame do developers apply when they select an algorithm?
- What frame do developers apply when they are interpreting the results?

Participants

This study involved Undergraduate students from the School of Computer Science, University of Birmingham. We accept that this might be seen as a limitation in the study but argue in our defence, that many implementations of 'routine' AI / ML will be performed by junior employees, i.e., recent graduates, in companies, and that their knowledge of methods could be similar to those of our student participants. The study employed in-depth qualitative investigation, involving interviews with three participants. The design of the study was approved by the University Ethics Review Board. Prospective participants were approached based on their experience in computer systems and software engineering, and the three who agreed to participate did so voluntarily.

The project

Rather than simply ask people to analyse datasets, we set them the challenge of manipulating the data to produce different levels of 'bias' in the output. In this way, they were working on a project that aimed to produce a 'bias machine' that could dial up or down 'bias' (from no bias to high levels of bias). Since the participants chose to address the health care system, working with medical datasets, 'bias' could be defined in terms of 'social' factors (age, gender, race, social class, level of healthcare etc.) or 'medical' factors (e.g., disease prevalence, predictive validity etc.). The activity began with a search for suitable datasets; the definition of suitability included factors such as size, representativeness and coverage of data, type of questions that could be addressed using these data etc. Most of the medical datasets have missing or incomplete data on patients, or certain patient groups or medical conditions could be over- or under-represented. Often this will require the analysts to modify the dataset, e.g., through normalisation, through consistency of labelling, through ensuring a balanced distribution of factors in the data. These processes could introduce distortions to the original dataset. Furthermore, processes of attribute sampling (in which elements in the dataset are selected on the basis of their relevance or expected contribution to the research question) can also introduce bias into the choice of data. In this project, two datasets were selected on the basis of record sampling, i.e., the project would use the complete dataset and there was little need for cleaning of the data (e.g., in terms of dealing with missing values):

- The Tobacco dataset to make predictions about future smoking behaviors relative to age and race.
- The vision and eye health dataset showed patients who were severely visually impaired, or blind based on a person's age, gender and race.

Working with datasets

The datasets were collected from a US website that provides access to a huge number of large datasets from different disciplines. Since the project focused on the medical field, the participants selected two databases to test bias using different AI / ML algorithms. The selected datasets were chosen for their size and variation. At first, participants kept 100% of the data so that they could demonstrate how selective filtering by the algorithms could force bias. Unfortunately, datasets were still subject to problems in sampling, labelling and representativeness. For example, certain race, gender or age groups were disproportionately represented which reflect the demographics of the source of the data but could have an impact on the quality of the output from the algorithms. However, decreasing the quantity of data for a particular group (perhaps to provide 'balance' in the dataset) by 6% was enough to produce a marked change in the results.

Working with algorithms

Having selected datasets, participants then explored different algorithms. Three different algorithms were applied (Neural Network, K-NN Classifier and Linear Regression). Linear Regression uses data points to identify a trend, and then makes a prediction by extrapolating this trend; K-NN Classifier predicts classes based on the percentage of given parameters; Neural Networks classify data based on two-thirds of the dataset for training, and one-third for testing. Participants applied the selected algorithms to the datasets, modifying parameters in the algorithms or datasets to introduce different outcomes. The aim was to identify parameters that would reliably produce 'biased' results (and to confirm that these parameters could also be specified to minimise bias). Meaning, some fields were removed and that affected the data, for instance, year end was removed (from tobacco dataset), so data collected at the end of a year (overlapping into the following year) would be treated the same as data collected at the start of the year. Moreover, Location would likely have a large effect on data - poorer areas with worse healthcare would have quite different results compared to richer areas. The sample size had also been ignored or in one dataset was seen as an 'equal' spread but it was not representative of the population distribution of the USA (where the dataset was collected from) so could be biased to a particular race. Therefore, the results of this exercise formed the basis of group presentations of the design and development activity. The choice of algorithm is considered in the analysis and results section.

Interviews

Three participants were interviewed on an individual basis to explore their beliefs and understanding of bias in the study. These interviews were conducted using the Critical Decision Method, with the probes shown in Table 1. Each interview took approximately 45-60 minutes.

Table 1: CDM questions

Probe type	Probe content (the interview question)
Cues:	Think about the last time you worked on machine learning (ML) project? What was about? Can you please talk in general about your project on Bias in Machine learning? What was the most difficult thing about the project? What were the feature that you were looking for in a dataset? What made a good dataset? What were the feature of the algorithms you apply? what made a good algorithm?
Information:	Why you think that the feature you defined first, were the best feature you would use? Why do you think this dataset is the best choice for your model? Can you explain the dataset? Do you think the data can be improved? How where the data collected? What are the reasons for selecting these specific datasets? Who would benefit from this?
Analogues & Experience:	What do you think is the main purpose of machine learning? Did you have previous experience with other algorithms? What about that previous experience did it seemed relevant to this project? Were you reminded of any previous experience using this method? What specific training or experience was necessary or helpful in making this project?
Standard Operating Procedures:	What is the process that you use to make these algorithms? (what data sets, where from, how validated, how sampled) Can you draw a flowchart or timeline for doing this? what decisions happen at each point… What is the process that you use to make the final decision? Can you draw a flowchart or timeline for doing this?
Goals:	Is there different output you can or want to make? Do you think the model can present more biased results? How? What do you think is the main purpose of presenting bias? Did the model present the specific goals you planed?
Assessment:	How do you know that you produced a biased result? What would make a good bias result? What would make a good unbiased result?
Mental model:	How do you train the models? How do you tune the algorithms? What model fitting do you do?
Decision making & Options:	Which algorithms do you think would be most useful for this application? Which ones did you use? Why did you select these? Which other (data) method do you think would be most useful for this project? Why did you select these methods? What other courses of action were considered or available to you?

Analysis and Results

Interviewing the participants in this study was conducted after their team had presented the results of their project (implementing bias Machine Learning algorithms on medical datasets). We started the interviews by asking the CDM questions (table 1) in order, encouraging the participant to express themselves as much as possible. However, in this section, we are highlighting only the most salient questions and answers.

Why do the project?

We asked the participants "What do you think is the main purpose of presenting bias?", this was to understand their motivation and how they viewed bias. P1 said, "The main purpose was essentially an educational tool to show how easy it is to have bias algorithms and more like how hard it is to have unbiased algorithms to the point of almost impossible." P2 said the project was aimed at showing that not every algorithm is suitable for any dataset, people can misuse them. You will only get a bias output if your inputs or process are biased. On the other hand, for P3 it was much harder to define bias P3 said "We did not know what unbiased would be. I suppose that is one of the challenges because it was hard to tell."

How to define bias?

Answers to the initial question give us an idea of how participants understood bias and how this was related to some wrong and questionable mathematical results. They linked bias to how well or bad the numbers looked in the datasets and how many error values the algorithms produce. This point of view was confirmed when we asked participants to describe the features that made a good dataset. P1 states, "For datasets, we were looking for kind of equal spread attributes, so we had a dataset which has race, gender and where people lived and we wanted it to be equal in terms of spread, I guess, so we wanted an equal amount of every race to be included to make sure it was not biased in any way or another in that kind of sets." Although, p1 just described a sample bias, P1 is also believing that if the data is statistically balanced then it is 'good'. However, sometimes 'balanced' data is not representative of the real-world population. Moreover, as P3 said "If it had one group that was more likely to smoke and we knew for a fact that was not true, that would be wrong. For example, we had a high percentage of the population smoking, and the algorithm assumed it was from the American Indian / Alaska native racial categories. This prediction is totally unfair…because, in fact, American Indian/Alaska Native were hugely overrepresented in the dataset as in reality, this group only makes up about 0.8% of the total population." This is not only a bias or fairness issue, but with a certain part of the population overrepresented in the dataset the result of the model or the prediction probably will be wrong. P2 said, "It is a hard question, but you just can tell bias when you see wrong values, or the results do not make sense, and there were overlapping and that cause bias because it is just not working, and it will not give right predictions".

Reason for choice of algorithms.

The teams used three algorithms to test bias (Neural Network, K-NN Classifier and Linear Regression) the reasons for selecting these were, because they had some experience with these algorithms, they were among the popular algorithms in ML and these algorithms gave a clear biased result according to the group. The participants accepted that the selection of algorithm was not based on absolute principles, but more like testing and exploring what works. P2 stated that the algorithms they used seemed to be very popular and P1 said "Honestly, we struggle to find a GOOD algorithm, and the ones we tested were very unreliable, and most of them had so many big issues that we really could not see them being used in the light of the application we tested", while P3 said "we tried the leaner regression, K-NN classifier and the Neural Network, and they were different from each other. Not so sure, but they were good enough." What was much more difficult for participants was to provide a coherent definition of what they meant by 'good'.

The outcome of the algorithm

We asked participants if the model produced results that fitted the goals of their project (which is to present bias) and two out of three agreed that they had the result they worked to achieve. P1 stated "yes, it showed how easy bias is added. Even if it not intended because before we start adding bias to things, we also tried to make an unbiased version of the algorithms and that was still biased" P2 also confirmed this point by saying "yes, this project can raise awareness about how people can be wrong in using certain algorithms with certain datasets. Just make people aware that high-end tools are not necessarily clever. If you get bad data, you will get the bad stuff out. It depends on what you feed your algorithm also what algorithm you are using as well." While p3 was expecting more obvious bias results and address some of the limitations that led to their output "The results were unexpected. I think we were imaging something that would have been achieved if we had a bigger dataset. So, I guess that is one of the main limitations (having small datasets) well, it was the largest we can find, but still, the model needed bigger." P3 continued by saying "We limited ourselves in the medical field where we could have much larger datasets, and it would have been useful. Mainly the bias we are finding was more down to the very limited size of the dataset rather than any other limitation regarding the data."

Conscious or unconscious bias?

Our last point to investigate was the possibility of producing the same result unconsciously, P2 said "I think our results can be produced unconsciously because some people are just not being aware that bias can be produced and then being naïve in which algorithm you chose for a certain dataset. I mean we did not even have to try to get this result. We all tried not to be biased in a way because everything was just falling"

Where does bias lie?

To highlight the reason for the bias or where it comes from, we asked the participants what caused this bias, algorithms or data? P1 said "In terms of who is most at fault, I think the model itself and the algorithms we used are very biased. They put their own biases on top of the dataset, well I do not think the datasets are perfect, but they are considered not too bad either. We did not find a lot of extra bias, but the algorithms would take this small amount of bias in the datasets and extrapolated it. In other words, the dataset is what causes the bias, but the algorithm allows it to make a big issue." P2 also agreed by saying "It is both. If you have a certain dataset you have to study it, and you did some tailor to select what algorithm you should use and vice versa, if you have an algorithm you have to tailor the dataset, so it worked the algorithm, so I would say." However, P3 had a very strong opinion that datasets are the cause of bias, "Definitely, it all depends on the data, if the dataset is not representative, then the output is not representative either. I think people produce these outputs because of bias dataset, and it is much difficult for the algorithms to be biased compared to the actual data you put in with"

The Frames

From figure 1 and our interviews, an initial observation is that the 'data' that forms the dataset for AI/ML is both 'data' and 'frame' for the developers, in that the choice of dataset (in terms of what it contains and the relationships it defines) creates the definition of the 'frame' (in terms of which algorithm to select, what questions could be asked, the dataset balances between factors etc.). To develop this further, we selected the elements of the frame based on the participants' answers to the detailed questions about the datasets and the algorithms. In this content analysis, we only used elements that were mentioned by 2 or more of the interviewees.

The first frame was the dataset frame, and it shows what was the main features of selecting the dataset from the participant's perspective as a team. As Figure 2 shows five features led them to select the Tobacco and the vision datasets, first, it had to be from the medical field, it should be big enough for the algorithm to execute, test and learn. Further, datasets must be diverse (has different entries to explore bias). It is also better to have an equal spread of attributes because there is no point exploring bias with a very bias dataset. Finally, it should contain lots of personal information, one reason for this feature is because personal information even if not explicitly used by the algorithm, affects the outcome. Another reason is lots of real-world companies use this info in their processes. Therefore, selecting datasets that got all these characteristics in their beliefs will affect the fairness of the model. According to P1, the data is a critical explanation for any biased system even if it was not the only reason.

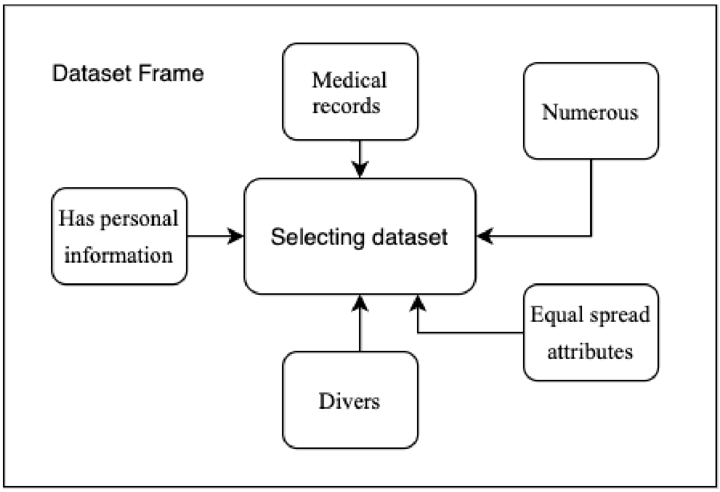


Figure2: dataset frame

The algorithm frame presents features of what makes a good algorithm and what are the reasons for selecting the three algorithms that were used in the project. It was most likely a straightforward process where lots of exploration was conducted to see what fits best. As Figure 3 shows, algorithms must be testable, used previously in ML, reliable and some members of the team (the participants) had an experience using the algorithm before the implementation. First let's make it clear that when the team decided to select an algorithm, they did not intend for a biased one, but the idea was to find a good algorithm to explore with the data they have. Unfortunately, even when not aiming for bias, the result suggests bias execution from the algorithms and unreliability, this was described earlier by the team and how they struggled to find a suitable algorithm, and how unreliable their algorithms were.

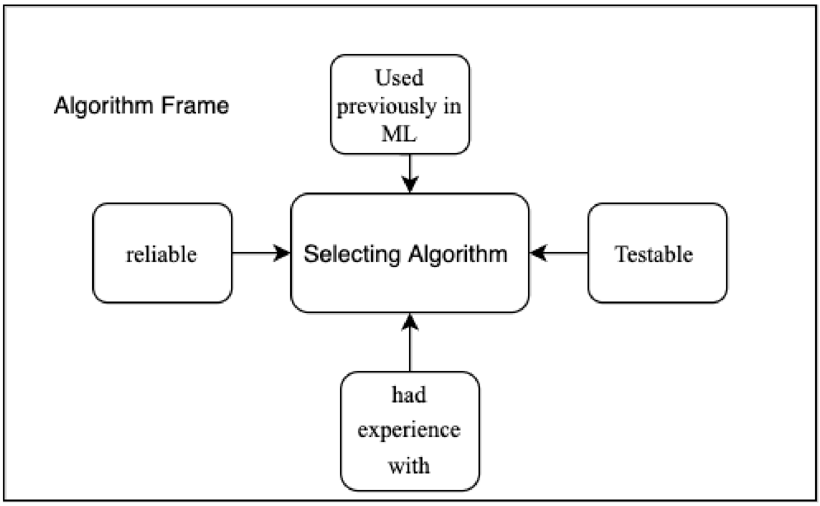


Figure3: algorithms frame

Figure 4 presents the initial goal of the model and how participants view bias. Figure 5 on the other hand, presents the final version of how participants interpret their results. These figures present different types of bias, and our argumentation in this comparison between these frames is that developers might not know the difference. This is because while working on the selection of the dataset and the algorithms, the questions that address the project is shifting when interacting with different process of cleaning data and testing algorithms. Seeing that developers often focus on Statistical, sampling, coverage bias (because they are using frames that address datasets and algorithms), at the expense of frames which address Epistemological bias.

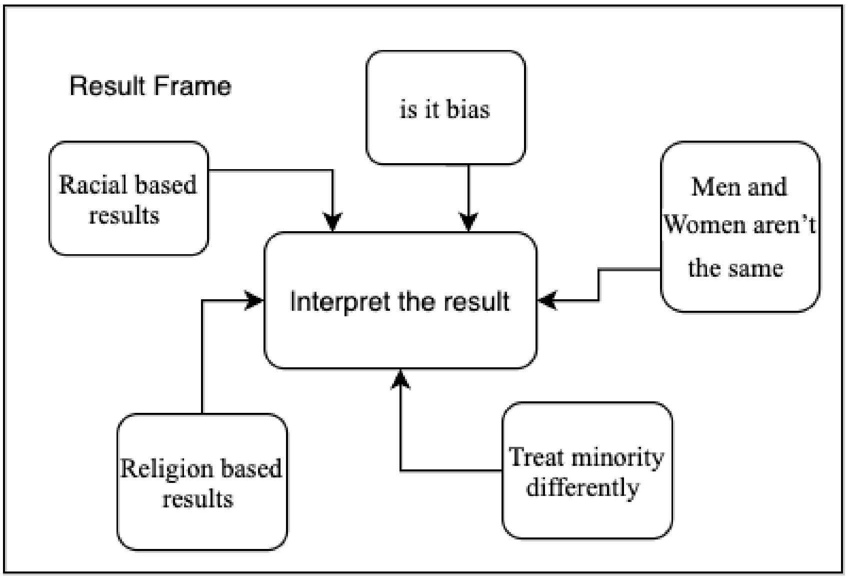


Figure 4: the initial result frame

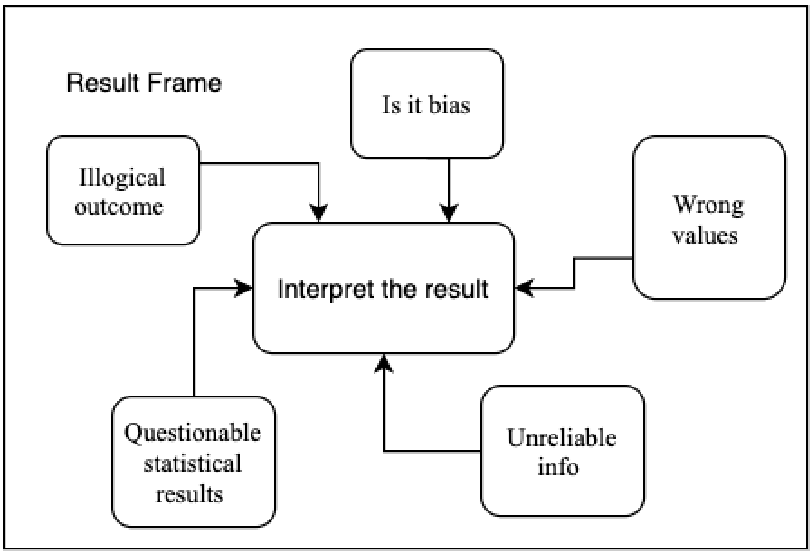


Figure 5: the final result frame 2

CONCLUSION

Conflict

We identified three frames (dataset frame, algorithm frame, and the resulting frame). When it comes to identifying bias in these frames, we believe that the data and the algorithms are usually considered by the developer, but it is not clear who would be responsible for the bias in the interpretation frame. Our results suggest that people change the initial version of the 'result frame' in the light of available datasets and algorithms without realising this. And this explains why we have two subframes for interpretation, the first one presents the 'expected result' at the beginning of the project which aims to present the ethical bias, and the second, a final result frame presents the actual outcome. This situation occurs when for example, the data need to be cleaned, adjusted and modified, resulting in differences between the original and prepared versions of the datasets. Moreover, the algorithm may need some data for training and some for testing, so this requires partitioning of the dataset. In other words, we end up with different conflicting interpretation frames. So, there is potential to introduce epistemological bias because the project changes slowly from project X to project Y (in that the original interpretation frame may have morphed into one that can be addressed using the revised dataset). Instead of project X, they have Project Y which has a nicely balanced dataset that executes well with the algorithm, but where the dataset might not reflect the real-world population, and the results did not address project X's problem.

These Frames presented three important components of any AI system (data, algorithms, developer), although. In figures (2-5) they seem to be different processes, they compete and complete each other in order to create a system. And when we say compete, we mean when you change the other components to fit one of them. For instance, when changing the original data to fit the algorithm, in this case, the algorithm has an important rule in the process. Oppositely, when the data is the critical one, every process to create the system will change according to the dataset.

Implications

This project shows how the developers of AI / ML might not take a narrow perspective on 'bias' (as a statistical problem rather than a social or ethical problem). At the end of the project, most of the participants' recommendations were based on technical implementation and awareness. This is not because they were unaware of these wider concerns but because the requirements relating to the management of data and the implementation of algorithms might narrow their focus into technical challenges. Consequently, bias outcomes can be produced unconsciously because developers are simply not attending to these broader concerns. There is a responsibility to think about bias, but it is not clear where in a work system that responsibility lays. We might think it is the responsibility of the programmer to consider bias when they build the program, but this study suggests that this is quite difficult for the programmer to think about social bias because they are too busy thinking about algorithms and data. The work system does not identify someone responsible to think about other types of bias. However, we can't always blame the programmer alone, we believe that someone else should be responsible for the result frame interpretation within the work system and find a way to link all these frames, so the bias in the result frame should be checked independently of the programmers and the programmers should be made to change what they were doing to minimize that bias. Therefore, creating accurate and effective model is important but so is ensuring that all races/ethnicities and socioeconomic levels are adequately represented in the data model (O'Neil 2016). Methods to debias machine learning algorithms are under development (Gianfrancesco et.al 2018) as are improvements in techniques to enhance fairness and reduce indirect prejudices that result from algorithm predictions (Kamishima et.al 2012).

REFERENCE:

Anjomshoae, S., Najjar, A., Calvaresi, D., & Främling, K. Explainable agents and robots: Results from a systematic.

Baber, C., McCormick, E. and Apperly, I., (2020), A framework for Explainable AI, *Contemporary Ergonomics 2020*

Baber, C., McCormick, E. and Apperly, I., (2021), A human-centered process model for Explainable AI, *Naturalistic Decision Making, 2021*

Borgo, R., Cashmore, M., & Magazzeni, D. (2018). Towards providing explanations for AI planner decisions. arXiv preprint arXiv:1810.06338.

High-Level Expert Group on ArtificialIntelligence (2019) 'High-Level Expert Group on Artificial Intelligence Set Up By the European Commission Ethics Guidelines for Trustworthy Ai', European Commission. Available at: https://ec.europa.eu/digital.

Hoffman, R. et al. (2018) Explaining Explanation, Part 4: A Deep Dive on Deep Nets, IEEE Intelligent Systems, 33(3), pp. 87–95.

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. JAMA internal medicine, 178(11), 1544-1547.

Kamishima, T., Akaho, S., Asoh, H., & Sakuma, J. (2012, September). Fairness-aware classifier with prejudice remover regularizer. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 35-50). Springer, Berlin, Heidelberg.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on systems, man, and cybernetics, 19(3), 462-472.

Klein, G., Moon, B. and Hoffman, R.R., (2006 a). Making sense of sensemaking 1: Alternative perspectives. IEEE intelligent systems, (4), pp.70-73.

Klein, G., Moon, B. and Hoffman, R.R., (2006 b). Making sense of sensemaking 2: A macrocognitive model. IEEE Intelligent systems, 21(5), pp.88-92.

Hoffman, R. R. (Ed.). (2007). Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making. Psychology Press.

Minsky, M. (1975). A Framework for Representing Knowledge, Reprinted in The Psychology of Computer Vision, P. Winston.

O'neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Thomas, J. B., Clark, S. M., & Gioia, D. A. (1993). Strategic sensemaking and organizational performance: Linkages among scanning, interpretation, action, and outcomes. Academy of Management journal, 36(2), 239-270.

Thordsen, M. L. (1991, September). A comparison of two tools for cognitive task analysis: Concept mapping and the critical decision method. In Proceedings of the Human Factors Society Annual Meeting (Vol. 35, No. 5, pp. 283-285). Sage CA: Los Angeles, CA: SAGE Publications.

A Human-Centered Process Model for Explainable AI

Chris BABER, Emily McCORMICK, Ian APPERLEY
University of Birmingham

ABSTRACT
With the increasing use of inscrutable Artificial Intelligence systems to support human decision makers, there has been much interest in what it means for these systems to provide 'explanation'. In this paper, the concern is with applying a simple formalism that can express a minimal set of features that can be used to define an explanation. It is argued that few contemporary AI systems support this minimal set. Advice is provided on how future developments in explainable AI systems could adhere to this minimal set.

KEYWORDS
XAI; Explainable AI; Explanation.

INTRODUCTION

Developments in Artificial Intelligence (AI) have demonstrated impressive performance, particularly on well-defined domains such as image processing or game playing. However, the techniques that are deployed can be opaque for the human user which raises the question of how AI systems can provide explanation (Neerincx et al., 2018; Rosenfeld and Richardson, 2019) and there is growing requirement for explainable AI (XAI) in Regulatory frameworks. Having said this, in 2017 Google's research chief, Peter Norvig, pointed out of the irony of expecting computers to provide 'explanations' when humans can be poor at doing this.

Much of the work on explainable AI (XAI) leans heavily on a computer-centric perspective (Springer, 2019). For example, Holzinger et al. (2020) assume that human and AI system have equal access to a 'ground truth'. From this, explainability *"…highlights decision relevant parts of machine representations…, i.e., parts which contributed to model accuracy in training or to a specific prediction."* In common with much of the XAI literature, this does not provide a role for the human, other than as passive recipient. The implication is that the AI system is able to introspect on its own processes to generate an explanation. The resulting explanation is then presented to the user, with description of the AI system's processes or the features ('decision relevant parts') that it has used. In this way, an explanation is simply a recommendation (from the AI system) plus the features that relate to this. As Miller (2017) notes, a problem with such an attitude is that it is based on the designer's intuition of what makes a 'good' explanation rather than on a sound understanding of how humans respond to, and make use of, explanation. This does not indicate why *some* features were selected or why the recommendation is appropriate to the user's concerns. Nor does it situate explanation in the wider organisation; it is likely that an explanation for the analyst will be distinct from that for the person managing data collection or the manager who will be briefed by the analyst.

For Holzinger et al. (2020) aspects of the situation (defined as a ground truth) are combined into a statement; that is, the explanation is simply an expression of this statement. This implies that there is a linear interpolation from features to explanation. This is similar to Hempel and Oppenheim's (1948) 'Covering Law Model' which was concerned with the ways in which Historians might explain an Event in terms of antecedent Causes. However, 'ground truth' (assumed by Holzinger's process model and by the covering law model) is seldom fully defined (leading to ambiguity in the selection of relevant features). This means that simply stating the situation aspects without an indication of why these (rather than other aspects) were selected might not lead to a useful or usable explanation.

Hoffman et al. (2018) provide a comprehensive review of literature relating to explanation. From this review, explanation involves sensemaking by the human (to contextualise the output of the AI system) and we agree that an appropriate framework for considering this is the Data-Frame model of sensemaking (Klein et al., 2007). Further, sensemaking (and its relationship with explanation) relies on the recognition that the process (of providing and receiving an explanation) must be reciprocal, iterative, and negotiated. This process relies on 'explainer' and 'explainee' reaching alignment. In other words, explanation involves 'common ground' (Clark, 1991) in which there is sufficient alignment in understanding for conversation to proceed. The nature of the conversation will depend on the situation in which the explanation is being provided and the goals of the explainee. For example, the explainee might be a 'trainee' who seeks to understand the explanation to learn criteria for a decision or might be an 'analyst' using the recommendation from the AI system to apply as a policy.

A PROCESS MODEL FOR EXPLANATION

Figure 1 illustrates the relationship between an Explainer and an Explainee in a Situation (Baber et al., 2020). A Situation has *features* which are analogous to the notion of 'data' in the Data-Frame Model. We use the term 'features' (rather than data) because the word 'data' has a narrow definition in the AI literature. Notice that in figure 1, relations are indicated by ≈ to indicate that these relations are partial, provisional and approximate and, just as the Data-Frame Model emphasises, require continual monitoring, checking and refinement. While the Data-Frame Model uses the term frame, this also has a privileged meaning in the AI literature. So, we adopt the term Relevance (Sperber and Wilson, 1982) to refer to the rationale for why features are selected by explainer or explainee. So, relevance could be defined by one or two features, F, or a cluster of features, C, or a belief, B (which allows predictions to be made about Features, Clusters and Situations), or a Policy, P (which associates Actions with the Situation). It is important to note that 'relevance' is relative, i.e., the definition of relevance would depend not simply on the features in the situation but on the prior experience of the people and their goals; the same situation could result in different Situation Models for the people experiencing it.

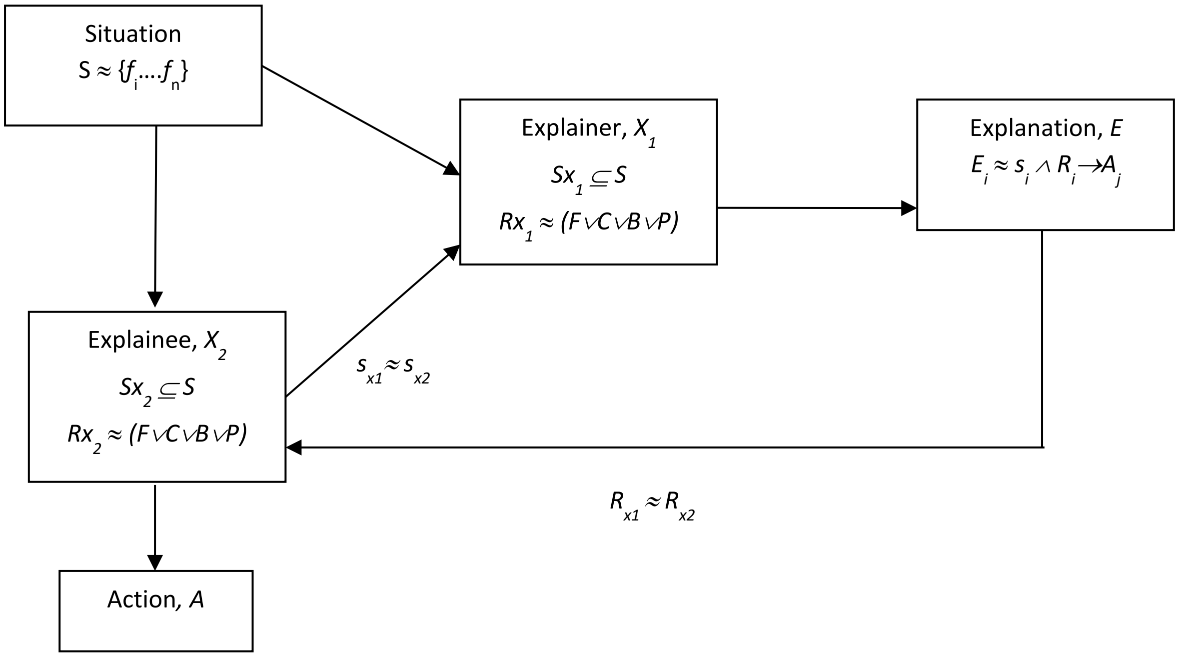


Figure 1. Defining Explanation

The Explainer and Explainee attend to features in a Situation (i.e., their 'situation model', *S*, consists of features which are selected as relevant to each person). From this, an explanation *could* involve overlap between 'situation models'. That is, $sx1 \approx sx2$. For models like that of Holzinger et al. (2018) the process of 'alignment' seems to be one-way (from computer to person) and assumes that this will result in the person adopting the same situation model as the computer. But, for a human-centered approach, this does not feel plausible. Rather, there needs to be scope for the alignment of situation models to be open to negotiation and dispute. We assume that explanations will involve second-order situation models, i.e., accounts which, at a later time, summarise the situation to another explainee. This means that, in order for 'situation models' to make sense to both parties, there needs to be alignment between definitions of 'relevance'. In this way, the primary means by which explanation operates is through alignment of 'situation model' *and* 'relevance' held by explainer and explainee, often (but not always) to encourage an Action on the part of explainee. However, we agree with Lipton (2016) that one must not treat 'explanation' as a monolithic concept. That is, different Situations (and different explainers and explainees) will require different types of explanation.

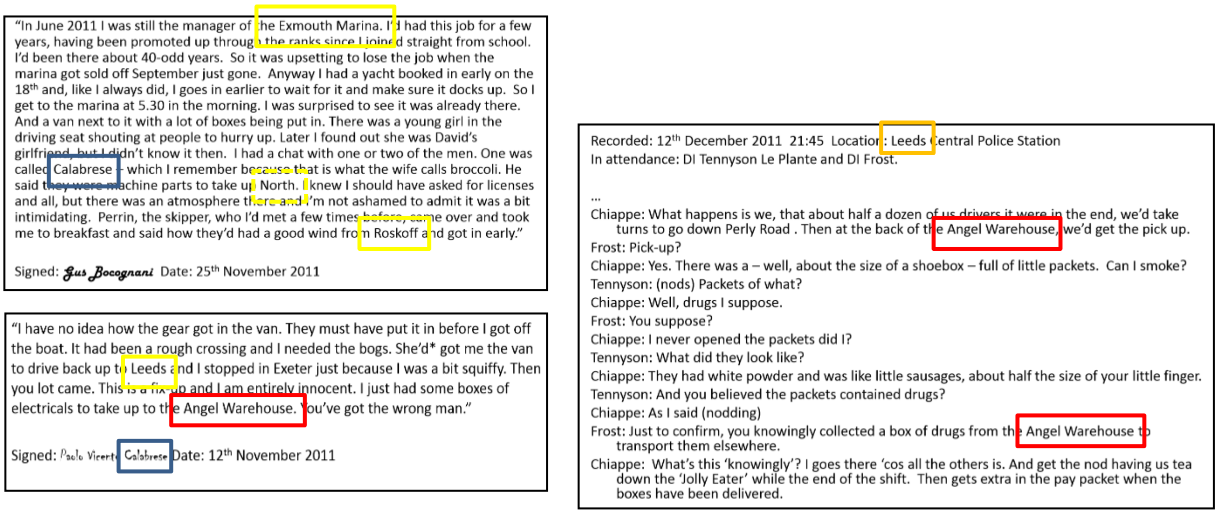
TYPES OF EXPLANATORY DISCOURSE

Implicit in the dichotomy, of situation model plus relevance, is the assumption that either of these can be 'surfaced' (i.e., brought to conscious awareness and expressed in words). Surfacing features of a Situation Model that are Relevant means not only an ability to introspect on our cognitive processes but also an ability to put the tacit knowledge that this implies into words. Further, AI system (particularly 'deep' AI) will be unable to introspect on its own processes. But explanations between humans seems to recognise this problem and we have techniques for managing explanatory discourse that enable us to challenge and explore this (Miller, 2017) – and these techniques have not been commonly applied to XAI (Miller et al., 2017).

We assume four types of explanatory discourse in which Situation Model or Relevance are Aligned or Challenged:

	Align	Challenge
Situation Model	Explainer draws attention to specific features in the situation.	Explainee disputes the indicated features and requires clarification of the situation model being applied.
Relevance	Explainer presents underlying rationale for the situation model.	Explainee offers alternative definitions of relevance or appeals to 'counter-factual' (or 'what if') examples, e.g., what if a given feature was present or absent.

To illustrate these types of explanatory discourse in human activity, figure 2 shows evidence used in the *North x South-West Exercise* for Intelligence Analysts (Baber et al., 2015, 2016). Highlighted features (within boxes) require knowledge of UK and Bretagne geography. Thus, to 'explain' the link between the pieces of evidence in figure 2, explainer and explainee need to agree on geography, e.g., the 'Angel Warehouse' is in Leeds, Leeds is 'up North', Exmouth and Leeds are connected by road, Roskoff and Exmouth are separated by an expanse of water.



"In June 2011 I was still the manager of the Exmouth Marina. I'd had this job for a few years, having been promoted up through the ranks since I joined straight from school. I'd been there about 40-odd years. So it was upsetting to lose the job when the marina got sold off September just gone. Anyway I had a yacht booked in early on the 18th and, like I always did, I goes in earlier to wait for it and make sure it docks up. So I get to the marina at 5.30 in the morning. I was surprised to see it was already there. And a van next to it with a lot of boxes being put in. There was a young girl in the driving seat shouting at people to hurry up. Later I found out she was David's girlfriend, but I didn't know it then. I had a chat with one or two of the men. One was called Calabrese – which I remember because that is what the wife calls broccoli. He said they were machine parts to take up North. I knew I should have asked for licenses and all, but there was an atmosphere there and I'm not ashamed to admit it was a bit intimidating. Perrin, the skipper, who I'd met a few times before, came over and took me to breakfast and said how they'd had a good wind from Roskoff and got in early."

Signed: *Gus Bocognani* Date: 25th November 2011

"I have no idea how the gear got in the van. They must have put it in before I got off the boat. It had been a rough crossing and I needed the bogs. She'd* got me the van to drive back up to Leeds and I stopped in Exeter just because I was a bit squiffy. Then you lot came. This is a fix-up and I am entirely innocent. I just had some boxes of electricals to take up to the Angel Warehouse. You've got the wrong man."

Signed: Paolo Vicente Calabrese Date: 12th November 2011

Recorded: 12th December 2011 21:45 Location: Leeds Central Police Station
In attendance: DI Tennyson Le Plante and DI Frost.

…
Chiappe: What happens is we, that about half a dozen of us drivers it were in the end, we'd take turns to go down Perly Road . Then at the back of the Angel Warehouse, we'd get the pick up.
Frost: Pick-up?
Chiappe: Yes. There was a – well, about the size of a shoebox – full of little packets. Can I smoke?
Tennyson: (nods) Packets of what?
Chiappe: Well, drugs I suppose.
Frost: You suppose?
Chiappe: I never opened the packets did I?
Tennyson: What did they look like?
Chiappe: They had white powder and was like little sausages, about half the size of your little finger.
Tennyson: And you believed the packets contained drugs?
Chiappe: As I said (nodding)
Frost: Just to confirm, you knowingly collected a box of drugs from the Angel Warehouse to transport them elsewhere.
Chiappe: What's this 'knowingly'? I goes there 'cos all the others is. And get the nod having us tea down the 'Jolly Eater' while the end of the shift. Then gets extra in the pay packet when the boxes have been delivered.

Figure 2. Items of Evidence for Investigation

As the first step in creating an explanation, Explainer and Explainee should attend to the same features , i.e., $S_{x1} \approx S_{x2}$. So, both people attend to the highlighted sections of figure 2 (if they do not, then the Explainer could point to each of these). But this is not sufficient to guarantee an explanation because the definition of relevance might differ between Explainer and Explainee, i.e., $R_{x1} \neq R_x$. From this, the Explainer wants to change the Explainee's notion of relevance so that it overlaps with part or all of the Explainer's notion of relevance, i.e., $\Delta R_{x2} \approx r_{x1} \subseteq R_{x1}$. Thus, the contents of the 'boxes' that are being transported by yacht and van, e.g., 'machine parts', 'electricals', or 'shoeboxes…of white powder', could be inferred by the Explainer to be the same thing; reference to 'electricals' or 'machine parts' could be deliberately misleading (based on the belief that the real content of the boxes is 'white powder'). The Explainer might point out ambiguity in the definition of 'contents' – in the expectation that the Explainee would recognise this. Or the Explainer might adjust the Explainee's relevance in order to have the Explainee perform an action, i.e., $\Delta R_2 \approx r_1 \subseteq R_1$ *and* $A_2 = \Delta s_2$. Believing that the 'boxes' contain 'white powder' (rather than electrical goods), the Explainer might seek to persuade the Explainee to conduct further analysis, e.g., collect Forensic reports from the boxes that have been recovered, or seek other instances where 'boxes' have been mentioned in interviews or reports, or speak to other people etc.

There will be situations in which the explainer and explainee are not able to reach alignment on the definition of relevance. For example, presentation in Court requires the Explainer to reconstruct the Situation in sufficient detail for the Explainee (in this case judge, jury or barristers) to appreciate (a) the selection of Features and (b) the Relevance of these features to the Situation. Dispute or disagreement could arise if the Explainee does not accept the features or their relevance, e.g., the explanation (of the content of the boxes) rests on the belief that they do not contain 'electricals' or 'machine parts' and that they do contain 'white powder'.

HOW DO AI SYSTEMS SUPPORT EXPLANATION?

Langley (2019) defines the operation of an agent capable of producing an explanation as:

> *"Given: Knowledge defining a space of possible solutions;*
> *Given: Criteria for evaluating candidate solutions;*
> *Given: An annotated search tree that includes solutions found for some reasoning task…;*
> *Given: A query about why a solution ranks above others;*
> *Produce: An explanation why the solution is preferable".*

In terms of our process model, the first two items in this list relate to the Situation Model that is used by the agent, and the second two relate to its definition of Relevance. From Langley's (2019) perspective, the purpose of 'explanation' will be to present the agent's situation model and relevance. But this seems to assume that alignment really means acceptance by the user. So, from this definition, explanation cannot be challenged. In other words, this definition rests on the assumption of transmission of the Explanation to the user rather than an explanatory discourse. Further, while the 'situation models' that humans create might be causal (e.g., in terms of plausible 'causes' of a given event or feature), it is more likely that the models that machines create are relational (e.g., correlation, regression, distance, similarity). This leads to the subtle problem of mistaking correlation for causation, i.e., the human could misinterpret correlations, on which the AI systems depend, for either causal (that is generalisable) relations or predictive beliefs. But neither of these (causal relations or predictions) are integral to the AI's algorithms. Next, we consider examples of how AI systems present explanations.

Features and Clusters

Many approaches to XAI require the user to infer the relevance of specific features to a recommendation. A popular approach to XAI involves explanation-by-simplification. For example, Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro et al., 2016) uses a specific instance which concentrates on local fidelity, i.e., the relations of that specific instance. In effect, the approach echoes the logic of the covering law.

Figure 3 shows alerts in financial trading. The implication is that this will allow the trader to ascertain the key features which led to an alert being raised. In this case, the explanation is the AI system's situation model. However, the display solely of features does not allow the trader to interrogate the underlying beliefs that led to the AI system raising an alert or to question its situation model. Moreover, the display is intended to motivate the user to conduct further investigation (probably drawing on other information sources) and, as such, cannot be, of itself, an explanation.

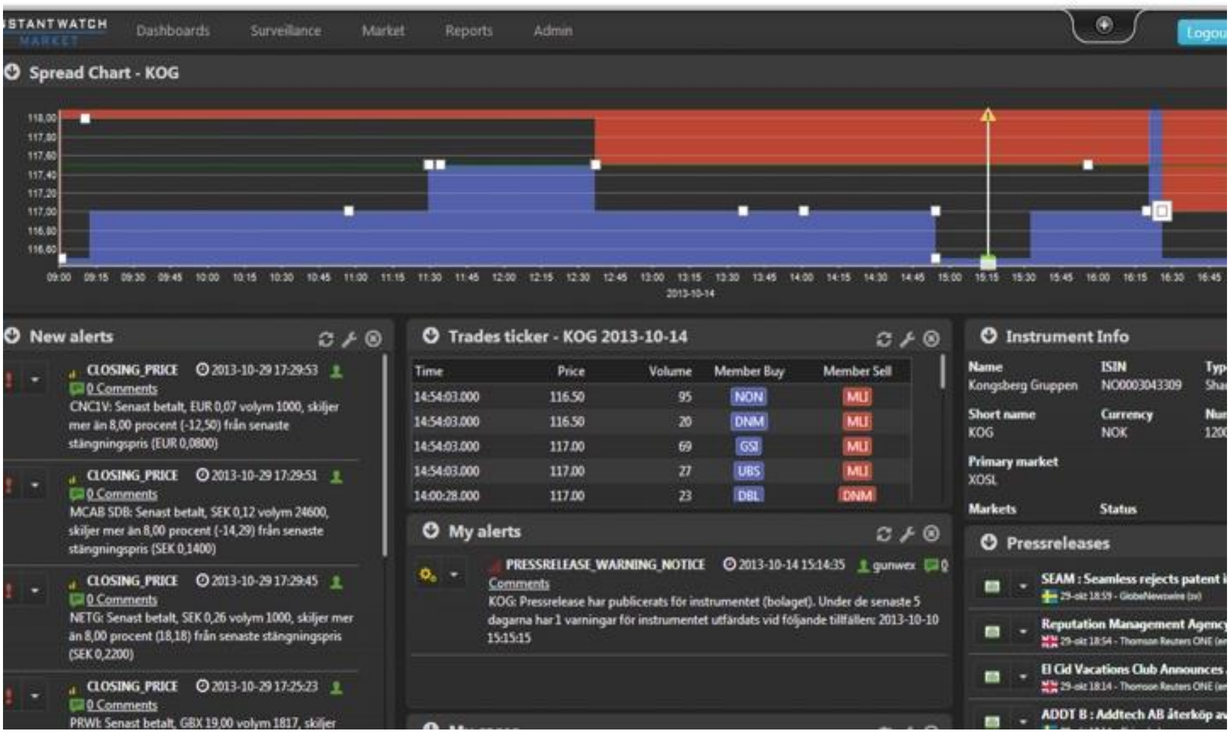


Figure 3: Monitoring stock market activity [https://www.trapets.com/services/instantwatch-market/]

Collating features into charts and tables provides the analyst with a summary that can be interpreted in terms of rules. Indeed, an experienced analyst might recognise recurring 'patterns' across different instances. That is, similar activities might produce displays that have similar visual appearance that the analyst can associate with particular activities. In this way the 'alert' relates not only to specific features but to the groupings of these features. While this might aid recognition-primed decision making (Klein, 1989) it does not provide access to the underling rules the AI system used to generate the clusters (and could result in the user either anthropomorphising these rules or making assumption about which rules could have been applied).

Beliefs

In figure 4, rules used to reach a loa decision are listed, together with an indication of whether the rules have been met or breached (with pass / fail, colour coding, accept / decline). In this way, the computer's rules are exposed to the human decision maker. The textual explanation, at the bottom of figure 4, is clear and concise. What is not apparent here is whether the user is able to apply counter-factuals to the decision. For example, if we consider the column for 'Application 4' in figure 4, the heuristic rule base identifies 'loan criteria, etc.' as below criteria, but what might happen if the applicant was able to amend this?

Decision Explanation Illustrator

	Application 1		Application 2		Application 3		Application 4		Application 5		Application 6	
Factual Rule Base												
Affordability test	FAIL		PASS		PASS		PASS		PASS		PASS	
Number and amount of bankruptcy	NONE		>=1		NONE		NONE		NONE		NONE	
Number of IVA & CCJ	NONE		NONE		NONE		NONE		NONE		NONE	
Number of payday loans	NONE		>=1		NONE		NONE		NONE		NONE	
Decision: Automated Application Acceptance or Decline	DECLINE		DECLINE		ACCEPT		ACCEPT		ACCEPT		ACCEPT	
Heuristic Rule Base	*F*	*R*	*F*	*R*	*F*	*R*	*F*	*R*	*F*	*R*	*F*	*R*
Unsecured loans	*Application declined in early stage*		*Application declined in early stage*		0.79	0.21	0.64	0.36	0.27	0.73	0.64	0.36
Secured loans					0.64	0.36	0.74	0.26	0.74	0.26	0.75	0.25
CCJ, IVA, Bankruptcy & payday loans					0.76	0.24	0.66	0.34	0.76	0.24	0.66	0.34
Searches					0.78	0.22	0.76	0.24	0.81	0.19	0.62	0.38
Credit score					0.79	0.21	0.61	0.39	0.62	0.38	0.73	0.27
Loan criteria, property valuation & property type					0.95	0.045	0.04	0.96	0.04	0.96	0.88	0.12
Predicted Output					0.96	0.04	0.07	0.93	0.07	0.93	0.89	0.11
Decision: Fund or Reject	REJECT		REJECT		FUND		REJECT		REJECT		FUND	
Textual explanation for a rejected application	Application has failed affordability test		Applicant have inadequate number and amount of bankruptcy and payday loans				-The property is poor and it has failed mortgage valuation. -The loan application do not fit our product-plan (loan criteria).		-The applicants have bad unsecured loan. -The property is poor and it has failed mortgage valuation. -The loan application do not fit our product-plan (loan criteria).			

Figure 4. Illustrating Beliefs in loan underwriting [Sachan et al., 2020]

Policy

In Deep (or Reinforcement) Learning, the AI system seeks to discover a Policy by which it can optimize reward (say, success in play a game) by performing Actions in specific situations. Accounts which reflect specific policy (in terms of the actions that AI systems take in response to situations) can be created as saliency maps (Greydanus et al., 2018). The saliency map can be used to infer the strategy that is being applied. While this need not reflect the policy (in terms of relationship between action and rewards that the agent is learning) it can allow the human analyst to form beliefs as to how the agent might behave in similar circumstances. However, it is not so easy to discern why the features were defined as Relevant, or even whether the AI system actually made use of these features. Combining a host of outputs, from the application of different algorithms, could allow the analyst to 'compare and contrast' the relevance of different features in terms of policy (figure 5). But this puts the onus on the user to infer an 'explanation' of the AI system's decision-making.

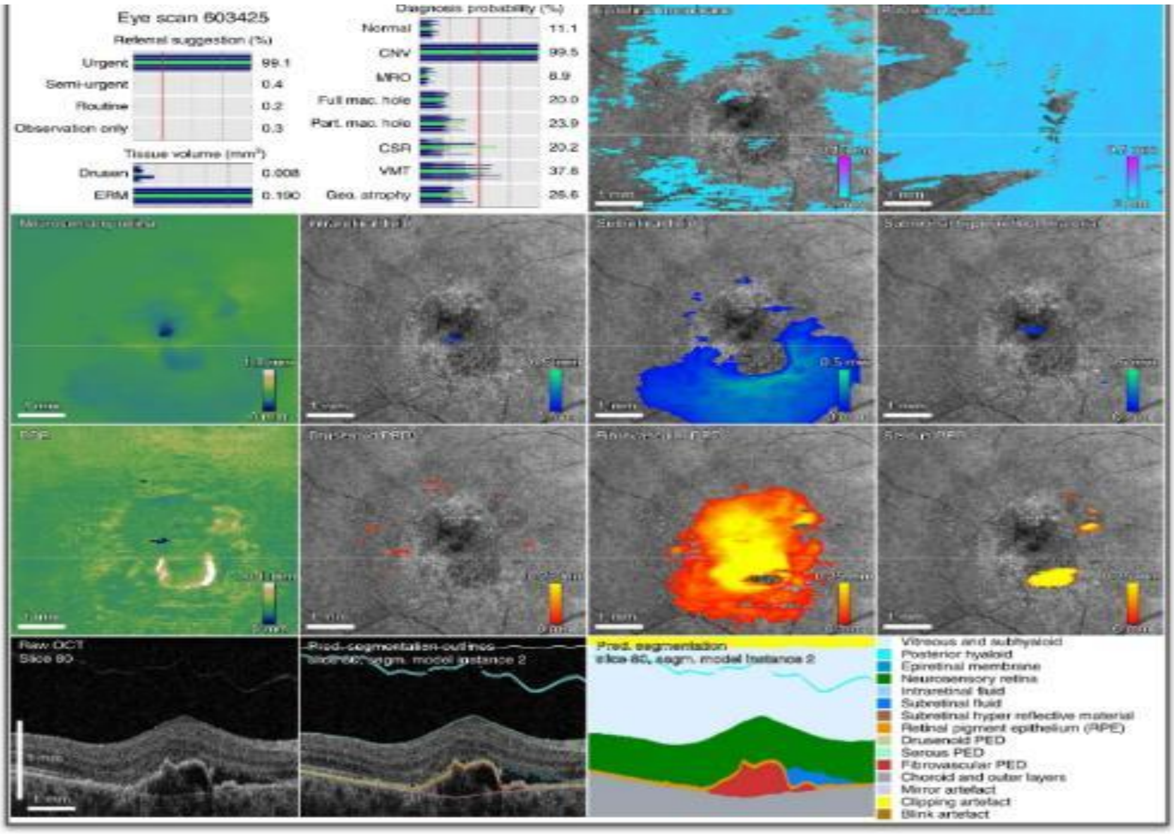


Figure 5. Multiple views of deep learning outputs for retinal diseases diagnosis [De Fauw et al., 2018]

Explanatory Discourse

In robotics, humans can ask questions of the robot that allow it to explain its reasoning. Fox et al's. (2017) eXplAInable Planning (XAIP) requires the robot to justify why it chose particular actions, etc. In terms of our notion of explanatory discourse (discussed earlier) this supports alignment of situation model (in terms of the features which are attended) or alignment of relevance (in terms of justification for an action). Recognising that human explanations can make use of choices of action in a situation, Borgo et al.'s (2018) developed XAI-PLAN. While the motivation for this, and similar work, seems to be the assumption that human explanations involve evidence and counter examples, it is not apparent that the ability to answer 'why' questions necessarily involves generation of a counter example. In other words, the issue of how the situation model or definition of relevance can be challenged has received less attention to date.

EXPLANATION AND DISAGREEMENT

In an experiment in which human and computer cooperate on an actions in a road traffic management task (Morar and Baber, 2017), the human participant needs to choose an action to manage a road network (defined by traffic volume and flow) and location on the map (figure 6). A computer provides suggestions as to which action to perform. Sometimes the computer is wrong. In this experiment, the Situation involves monitoring road traffic to reduce congestion (by altering traffic rate, through control using traffic lights, to reduce density). The user is provided with information on the identification of a ramp to join or leave a highway (shown as the highlighted box in the 'ramp metering' window), the location of the ramp (shown on the map), and the state of the road network at that location (shown by the bubble chart of traffic density and rate). To read the bubble chart, the following heuristics are applicable: Low density, Low rate: no response; High density, High rate: reduce rate; High density, Low rate: increase rate; Low density, High rate: no response.

In the 'ramp metering control' window (bottom right, figure 6), the user indicates which action to take, a reason for this action, and their interpretation of the bubble chart. Below this, the computer provides its solution. Users compare their responses with that of the computer, decide whether they wish to alter their response, and then use the 'submit' button to confirm this response.

Figure 6: User Interface for Road Traffic Task

Table 1 translates the requirements for this experiment into the terminology of our process model, to illustrate how concepts of relevance apply to different aspects of the experiment and how these might be aligned (by the human) in their decision making.

Table 1. *Mapping the Road Traffic Decision Task to the Process Model*

Element	Situation Model	Relevance	Alignment
$S_{x1} \approx S_{x2}$	Traffic on map; Behaviour of traffic on graph; ramp being monitored	Feature: ramp id / location Cluster: bubbles on graph Belief: definition of congestion	agree Features
$R_{x1} \neq R_{x2}$	Radio-buttons selected in Ramp Metering Control panel	Belief: Mismatch between Computer and User answers	agree Belief
$\Delta R_{x2} \approx r_{x1} \subseteq R_{x1}$	Selection of radio-buttons in situation model should agree	Belief: Computer is correct	agree Belief
$A_2 = \Delta s_2$	User Acts to change the Situation	Policy: Press 'Submit' button'	User is responsible for decision

An observation from Morar and Baber (2017) is that, rather than the computer supporting the users' decisions, there was often an assumption that the 'computer' provided information requiring confirmation. This is especially problematic if 'computer' reliability is "quite high", i.e., >80%, because it requires scrutiny of the recommendation at a level of detail that is not required if reliability is 'low' or 'perfect'. The Situation model uses Features attended by human and computer and a common user error was to miss differences relating to ramp id. In terms of Relevance there was a need to align Beliefs about the Situation, and a common user error was not to recognise computer misinterpretation of the graph, especially when this compounded ramp id errors.

CONCLUSION

By way of conclusions, we offer some basic guidelines for the implementation of XAI:

1. Explanation should be related to beliefs about the relationship between features that can directly affect the situation being explained (situation model), or can explain the majority of the situation (explanatory power), and are plausible (construct validity);
2. The Explanation should relate the goals of the explainer and explainee.
3. The explanation to suit the explainee's definition of relevance;
4. Explanations should be interactive and involve the explainee in the explanation;
5. Explanations should be (where necessary) actionable. The explainee should be given information that can be used to perform and/or improve future actions and behaviours;
6. There should be clarity in the definition of relevance used in the explanation: Define clusters (i.e., statistical model), belief (i.e., causal model) and policy (i.e., implications for action);
7. Explanatory discourse should allow challenge and the use of counter-examples to test the situation models and definitions of relevance that are employed in the explanation.

ACKNOWLEDGMENTS

The ideas developed in this paper have arisen from a study on ESRC Centre for Research and Evidence on Security Threats: CREST Project 277: Human Engagement Through Artificial / Augmented Intelligence.

REFERENCES

Amir, D. and Amir, O. (2018) Highlights: summarizing agent behavior to people, *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, 1168–1176.

Baber, C., Conway, G., Attfield, S., Rooney, C., Kodagoda, N., and Waller, R. (2015) How military intelligence personnel collaborate on a sense-making exercise, *12th International Naturalistic Decision Making Conference*, McLean, VA.

Baber, C., Attfield, S., Conway, G., Rooney, C. and Kodagoda, N. (2016) Collaborative sensemaking during Intelligence Analysis exercises, *International Journal of Human Computer Studies, 86,* 94-108.

Baber, C., McCormick, E. and Apperly, I. (2021) A framework for explainable AI, *Contemporary Ergonomics 2021*.

Borgo, R., Cashmore, M. and Magazzeni, D. (2018) Towards providing explanations for AI planner decisions. *arXiv*:1810.06338.

Clark, H.H. (1991) Using Language, Cambridge: Cambridge University Press.

De Fauw, J. et al., 2018, Clinically applicable deep learning for diagnosis and referral in retinal disease, *Nature Medicine*, *24*, 1342–1350

Fox, M., Long, D. and Magazzeni, D. (2017) Explainable planning, *arXiv*:1709.10256

Greydanus, S., Koul, A., Dodge, J. and Fern, A. (2018) Visualizing and understanding Atari agents, *arXiv*:1711.00138v5

Hempel, C. G., and Oppenheim, P. (1948) Studies in the Logic of Explanation, *Philosophy of science, 15*, 135-175.

Hoffman, R., Miller, T., Mueller, S. T., Klein, G., and Clancey, W. J. (2018) Explaining explanation, part 4: a deep dive on deep nets, IEEE Intelligent Systems, 33(3), 87-95.

Holzinger, A., Carrington, A. and Müller, H. (2020) Measuring the Quality of Explanations: the Systems Causability Scale (SCS), Künstliche Intelligenz, 34, 193-198

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H. and Wortman Vaughan, J. (2020) Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-14.

Klein, G.A. (1989) Recognition-primed decisions, In W. B. Rouse (ed.) *Advances in Man-Machine Systems Research*, *volume 5*, Greenwich, CT: JAI Press, 47-92.

Klein, G., Phillips, J.K., Rall, E.L. and Peluso, D.A. (2007) A data-frame theory of sensemaking, *Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making*, New York, NY: Lawrence Erlbaum Assoc Inc., 113-155.

Langley, P., (2019) Varieties of Explainable Agency, *ICAPS Workshop on Explainable AI Planning*.

Leite, R.A., Gschwandtner, T., Miksch, S., Kriglstein, S., Pohl, M., Gstrein, E. and Kuntner, J., 2017. Eva: Visual analytics to identify fraudulent events. *IEEE transactions on visualization and computer graphics*, *24*, 330-339.

Lipton, Z.C. (2016) The mythos of model interpretability, *arXiv*:1606.03490.

Miller, T. (2017) Explanation in Artificial Intelligence: insights from the Social Sciences, *arXiv*:1706.07269.

Miller, T., Howe, P., and Sonenberg, L. (2017) Explainable AI: beware of inmates running the asylum, *Proceedings IJCAI-17 Workshop on Explainable Artificial Intelligence (XAI)*

Morar, N. and Baber, C. (2017) Joint human-automation decision making in road traffic management, *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting*, Santa Monica, CA: HFES, 385-389.

Neerincx, M.A., van der Waa, J., Kaptein, F. and van Diggelen, J. (2018) Using perceptual and cognitive explanations for enhanced human-agent team performance, In Harris, D. (ed.) *EPCE 2018,* LNCS (LNAI), 10906, Springer, Cham, 204–214.

Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1135-1144.

Rosenfeld, A., and Richardson, A. (2019) Explainability in human–agent systems, *Autonomous Agents and Multi-Agent Systems*, *33*, 673-705.

Sachan, S., Yang, J.B., Xu, D.L., Benavides, D.E. and Li, Y. (2020) An explainable AI decision-support-system to automate loan underwriting, *Expert Systems with Applications*, *144*.

Sperber, D. and Wilson, D. (1986) *Relevance: Communication and Cognition*, Cambridge, MA: Harvard University Press.

Springer, A. (2019) Enabling effective transparency: Towards user-centric intelligent systems, *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 543–544.

How Experts Know That Their Intuition is Right? Naturalistic Study on Executive Search Consultants

Matylda GERBER
Warsaw School of Economics

ABSTRACT
Previous studies on the role of intuition in experts' decision-making allowed to discover strategy experts use to make correct decisions in situations in the absence of clear facts. This study aims to discover how expert executive search consultants know when to trust their intuition and how intuition influences their process of reasoning and decision-making. Six executive search consultants from Warsaw and London took part in the study (average experience = 13 years). Results of the study indicated that consultants use intuition to provide an initial appraisal of a candidate. Intuitive answer was found to be accompanied with a Feeling of Rightness (FOR) whose intensity was prompting the level of confidence in one's intuition. Based on results, a decision-making model was created that depicts how FOR influences consultants' process of reasoning and decision-making.

KEYWORDS
Intuition, Decision-Making, Experts, Executive Search Consultants, Naturalistic Study

INTRODUCTION

Previous studies on the role of intuition in experts' decision-making were conducted on police criminal intelligence analysts (Gerber, Wong, Kodagoda, 2016) and medical doctors. Results of the studies allowed to discover a strategy experts use to make correct decisions in situations in the absence of clear facts. The studied experts were using intuition to identify relevant problems and area for solution search. Afterwards, they were interpreting their intuition, what was allowing them to set a direction for data collection and analysis. While the process of data collection and analysis, they were gaining insight with a solution to the problem. Results of the studies indicated that experts use unsupported intuition to initiate the process of fining solution that could be justified with previously collected data. The purpose of the research on the role of intuition in experts' decision-making is to collect relevant knowledge which will allow to design a decision support system for intuitive decision-making. Such system would be helpful for situations where intuition is the only support because of lack of relevant cues based on which a decision could be made. The aim of the study presented in this paper is to discover what makes experts know that their intuition should be trusted. The study has an exploratory form and it is conducted on expert executive search consultants.

In recent years a search for talented leaders has become a priority for organizations (Guthridge, Komm, Lawson, 2008). The complexity of this process has led to the outsourcing of talent recruitment and selection, resulting in sharing responsibility of finding the right candidates with third parties - known as Executive Recruitment Firms or Headhunters (Greer, Youngblood, Gray, 1999).

Executive search consultancy refers to the process of attracting and selecting candidates through direct and personal contact by a specialist consultant who acts as an intermediary between the client and the candidates for the available position (Britton, Wright, Ball, 2000). The difficulty of executive search consultants' job is not only to correctly asses candidates' potential, but also to convince the client that the candidate is suitable for a searched position. Due to information asymmetries between a candidate and a consultant, executive search consultants need to make decisions based on incomplete information.

Clerkin and Lee (2010) highlight the scarcity of research on the executive search process. There are theories that frame executive search consultancy as outsourcing (Finlay, Coverdill, 1999), explain why young companies take over talented executives from competitors (Rao, Drazin, 2002) or clarify the expectation gaps between consultants and candidates (Britton, Wright, Ball, 2000). However, there is no research showing how executive search consultants deal with situations in the absence of clear facts in order to assess correctly candidates' potential and to find suitable candidates for a searched position.

There are studies showing that executive search consultants claim to rely on intuition in order to correctly evaluate candidates (Miles, Sadler-Smith, 2014). However, little is known about the way they use intuition to deal with such decisions and about the mechanisms that make them trust their intuition. As consultants need to justify their decisions to their clients, they cannot make decisions solely on intuition. That is why, the process which needs to be investigated is complex decision-making, where intuition goes in interaction with analytical reasoning.

There is a research field that aims to discover how an interaction between intuitive and rational reasoning unfolds. It is called Dual Process Theories (DPT) and it provides the most popular explanations about the way an interaction between intuitive and rational thinking allow us to think and to make decisions. Those theories assume existence of System 1 – autonomous, fast processing and System 2 – a much slower process, which requires working memory engagement (De Neys, 2017). The most popular point of view on the interaction between System 1 and System 2 assumes that people receive an automatic response from System 1 right after approaching a given problem. It is a default system. After a disclosure of a response from the System 1, the System 2 process is engaged. It either utilizes the automatic response or it introduces some changes. The System 2 intervenes, when it detects a conflict in reasoning (Evans, 2008).

Dual Process Theories are usually tested in the form of laboratory experiments. One of the most popular tools used to study the assumptions of DPT is Cognitive Reflection Test (Frederic, 2005). This test is composed of text tasks, such as:

A bat and ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

Result of studies indicate that study participants usually answer the question based on the result of System 1 processes, providing the answer to be $0,1. This result comes to mind automatically, even though the right answer is $0,05. Depending on time availability, intelligence and motivation, some people correct the intuitive response through System 2 engagement (Evans, 2011). The further studies within this concept indicate that there are also participants who are able to provide the correct answer immediately. This observation motivated the researchers to ascertain that System 1 processes could be responsible also for logical thinking, proving that different forms of intuition exist (De Neys, 2012).

As Dual Process Theories are predominately tested in form of laboratory experiments, there started to occur a discussion about the need to verify functionality of DPTs in natural decision-making environments. It will allow to create decision-making models for particular situations, which will enable to verify the functionality of Dual Process Theories in real decision-making situations (Evans, 2017).

The key matter to be solved around Dual Process Theories concerns the incentives that cause System 2 to intervene reasoning and decision-making process. A number of variables has been already proposed to determine System 2 occurrence. It is, among others:

- Individual abilities of a reasoner, such as cognitive capacity (Stanovich, 1999),
- The characteristic of decision-making environment i.e. time available to complete the task (Finucane, et al., 2000) or the characteristic of given instructions (Newstead, et al., 1992).

However, an additional concept has been recently introduced, postulating the existence of internal stimuli, responsible for triggering System 2 intervention. These stimuli are metacognitive experiences that are attached to System 1 processes. It is found that the result of a reasoning is conditioned not only by the information retrieved by System 1 and analyzed by System 2, but also by a second-order judgment, called meta-reasoning.

"Meta-Reasoning refers to the processes that monitor the progress of our reasoning and problem-solving activities and regulate the time and effort devoted to them. Monitoring processes are usually experienced as feelings of certainty or uncertainty about how well a process has, or will, unfold." (Ackerman, Thompson, 2017, p.1)

The meta-reasoning is based on feelings attached to intuitive cues. Those cues are not always reliable, but they allow to regulate our mental effort. In order to understand meta-reasoning, two levels of processes should be considered. The first level concerns object-level processes. These processes are responsible for carrying out cognitive work such as; perceiving, classifying, remembering and deciding. The second level operates on the meta-level which monitor the first level processes. The role of the second-level processes is to asses functioning of the first-level processes and to efficiently locate resources. The metacognitive monitoring indicates different levels of certainty and uncertainty informing if a set of processes lead to the right direction. The operation of metacognitive processes is visible in such situations as when high confidence in our answer motivates us to act and hesitation causes us to seek for more information or to change decision-making path. We may also feel unable to pursue a task, what causes us to seek help or to stop taking any further action. One of the meta-reasoning processes is

Feeling of Rightness (FOR) which is perceived to be attached to automatic – intuitive responses. FOR is found to inform to what extent decision-makers might trust in their intuitive responses (Thompson, Turner, Pennycook, 2011).

There is a number of studies indicating the role of meta-reasoning in the process of learning, remembering and comprehension. However, little is known about its role in complex processes such as reasoning and problem-solving (Ackerman, Thompson, 2017). That is why, the aim of the research presented in this paper is to verify how intuition and attached to it Feeling of Rightness influences expert executive search consultants' process of reasoning and decision-making. The study is provided on members of organization with the highest level of expertise. It is because, this skill-level is perceived to enable decision-makers to use intuition the most reliably (Dreyfus, 2004).

METHOD

Participants

Six executive search consultants from Warsaw, Poland and London, UK took part in the study. Their average work experience was 13 years. The participants were chosen with help of an executive search consultant who has a good reputation among clients and other consultants. The criteria for choosing participants was over 9 years of experience in executive search and a good reputation among clients and other consultants. From 32 sent invitations to take part in the study, 8 consultants agreed to meet and 6 of them agreed to participate in the study. The participants represent three different, international executive search companies, which have similar size and compete in the same sector.

Procedure

Appointments with study participants were set up at the time consultants had meetings with candidates for particular projects. The study consisted of three elements:

1. An interview with a consultant before meeting with a candidate.
2. Observation of the meeting with a candidate.
3. An interview with the consultant after the meeting with a candidate.

The three phases of the study allowed to capture how the decision-making process unfolds before, during and after the meeting with a candidate. The first phase - the interview before the meeting aimed to provide an understanding about key decision elements the consultant predicts to encounter. This phase allowed the researcher to capture elements she needs to focus on during the meeting with a candidate. The interview had semi-structured form. The consultant had to make introduction to the case, explain requirements a candidate needs to fulfill, describe strategies he/she plans to utilize in order to acquire relevant information and to describe assumed weak points of the candidate. As the study had exploratory form, the rest of questions were arising while the interview. The second phase – observation of the meeting with a candidate, allowed to see how the studied consultant verify predefined topics. The third phase – the interview after the meeting allowed to get understanding about consultant's reasoning process during the meeting with a candidate. This interview had unstructured form and the questions were asked on the basis of observed interesting elements while the meeting with a candidate.

Materials for analysis consisted of recordings of the interviews before and after meetings as well as notes made during meetings with candidates. The recordings were transcribed, and they became a base for analysis.

Data Analysis

A qualitative data analysis approach called the Emergent Themes Analysis (Wong, 2004) was used on transcripts of six in-depth interviews with executive search consultants from Warsaw and London. In order to get a general understanding of consultants' decision-making process, analysis tables were created. The tables were formed separately for decision-making process before and after the meeting and afterwards information from both tables were compared to each other. The analysis table for decision-making process before the meeting contain following information:

1. Requirements a candidate needs to fulfil (hard requirements, soft requirements, requirements which unfulfilled disqualify a candidate),
2. Intuition occurrence,
3. Consultant's individual strategies utilized during a meeting,
4. Manifestation of expertise,
5. Things that distinguishes a candidate from others,
6. Assumed weak points of a candidate,
7. Candidate's appraisal before the meeting.

The analysis table about the decision-making process after the meeting contains the same information as the previous table but additionally it reflects the following information:

1. First impression while meeting a candidate,
2. Candidate's appraisal after the meeting,
3. Rationalization process of unfulfilled hard requirements.

All the information in the tables were supported with excerpts from interviews.

RESULTS

In order to find the role of intuition in executive consultants' decision-making, the information from two tables (before and after the meeting) were compared and the following aspects were chosen to be captured:

a. How perception of a candidate changes before and after meeting?
b. How initial intuition towards a candidate influences the process of data collection and analysis?
c. How initial intuition towards a candidate influences interpretation of received information?
d. How experts' experience affects their understanding of situations?

After the first part of analysis, an interesting relationship was identified; depending on the first impression made by a candidate, the consultants were collecting and interpreting information in a different way. It means, in situations when the first impression was significantly positive, the consultants focused more on collecting supporting than rejecting information. Moreover, they tended to explain away contradictory information and to search workarounds in situations where a candidate did not fulfill important requirements:

[before meeting the candidate] 135. P1: We need a person who has experience in the sale of a large number of products [...] meaning we are interested in someone who already has experience in a large number of product types, and this is very important. I will just ask him who he was and what he sold, and if it turns out that the candidate sells something only to large factories, so-called key accounts or to distributors who later sells it further, then this is not our candidate.

[after meeting the candidate] 11. P1: He does not meet one of the key conditions, he does not manage a direct team. However, the character of his team is so close that, thanks to the fact that he has a number of other advantages, I think that we can go over the agenda.

Similarly, but in reverse, it looked in situations where there was a strong negative first impression.

316. P1: Sometimes it happens to me and it's the most often confirmed, although I spend an hour with a human being anyway, because I am a technical engineer, what I think after 30 seconds, it's cool that it seems to me but you have to check it out but it's already scary for me, that the candidate will get up, give his hand and say 'good morning' and I already know 'damn, hour wasted, nothing will come of it'.

In situations where a consultant did not have any strong first impression his/her form of collecting information wasn't focused neither on supporting nor on denying candidate's application. Consultants were collecting similar amount of positive and negative information concerning the candidate and they were generating possible positive and negative interpretations of situations.

[after meeting the candidate] 305. P6: Such a calm man. I: And you had it in green, or in red, or so in orange? P6: No, as a question mark. Yes, as a question mark.

[after meeting the candidate] 110. P6: I still do not know if he will be good. At the moment I can see that he fits my profile, that is, he worked in a similar business, he works in a loan business, he was sales manager, he was the managing partner, i.e. I see here sales and management skills.

Due to the fact that the analyzed first impressions were supported with collected information, it should be assumed that the initial feeling towards the candidate is not just a feeling of liking/not linking a candidate, but it is rather a more complex assessment of a candidate that is derived from many years of experience in recruiting candidates.

243. P2: [At an earlier stage of my career] I wasn't able to say so quickly that this is a good candidate. Having at the back of head all these previous candidates, whom I have seen in similar positions over the years. She takes position in my head yes or no immediately.

This initial feeling-based assessment of candidates is compatible with the concept of meta-reasoning, which indicates that the initial feeling experienced by decision-makers reflects information that there are implicit structures in decision-makers mind that fit well to the currently experienced decision situations. In case of executive search consultants, it could mean that a given candidate highly resembles candidates successfully accepted for a similar position.

Experts' Decision-Making Model

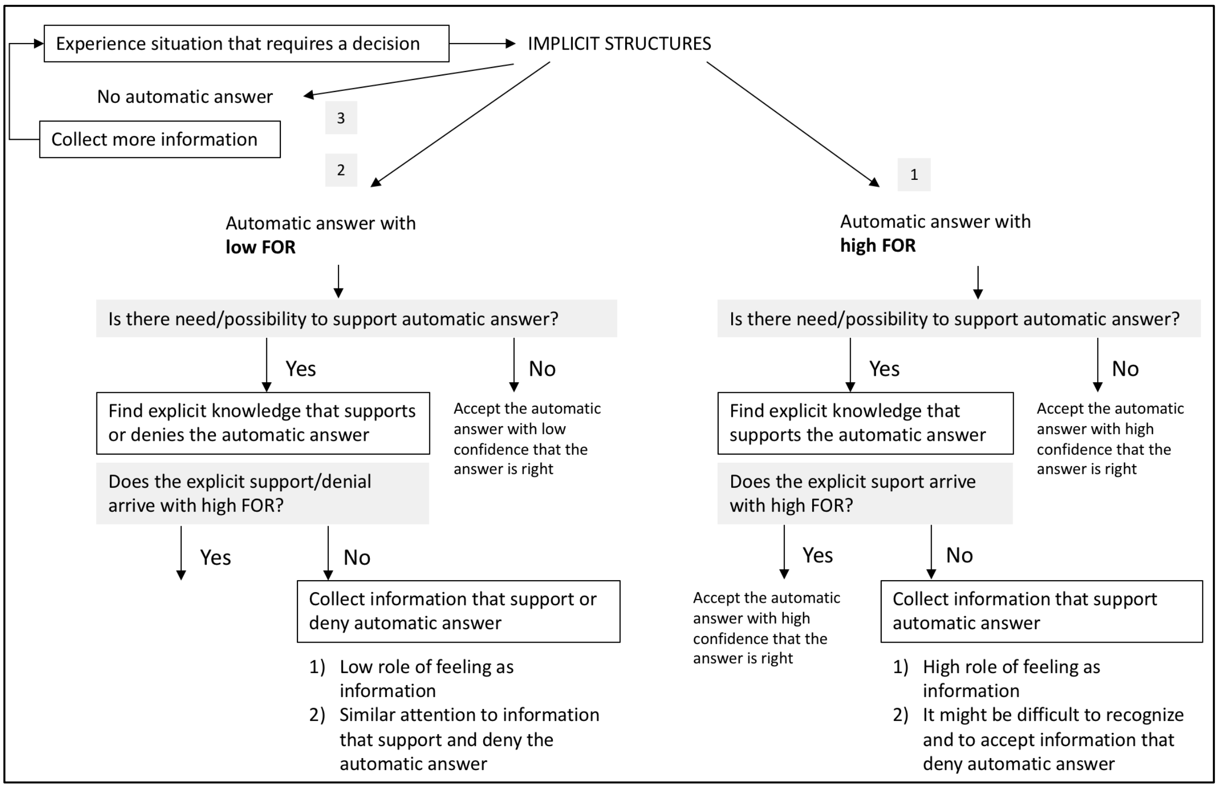
Five out of six executive search consultants were found to structure their reasoning and decision-making process in accordance to their initial intuition. It could be assumed that the intuitions were assisted with high Feeling of

Rightness, because they were highly convinced that their initial intuition was right. One consultant had automatic answer with low FOR and he doubted in accuracy of his intuition.

Based on analyzed decision-making processes of six expert executive search consultants and on the concept of meta-reasoning, a decision-making model was created. The model indicates how experts construct their process of data collection and decision-making, depending on the intuitive-automatic answer occurrence and a corresponding Feeling of Rightness. The model assumes three possibilities:

a. Automatic answer occurrence (i.e. the candidate fits to the given position) with high FOR,
b. Automatic answer occurrence (i.e. the candidate fits to the given position) with low FOR,
c. No automatic answer occurrence.

Figure 1. Experts' Decision Making Model – Executive Search Consultants



Source: Own Elaboration

Decision-making path 1: automatic answer occurrence with high FOR.
The first decision-making path in the model describes a situation when an automatic answer occurs with high FOR. High FOR indicates that there are implicit structures that fit well to the given decision situation. In the case of executive search consultants, it could mean that the candidate highly resembles successfully accepted candidates for similar positions. If there is no need or possibility to support automatic answer, decision makers accept the answer with a high confidence that the answer is right. High confidence occurs, because the strong positive feeling prevents from doubts occurrence. Executive search consultants' job specification requires explanation of their decisions, that is why, their decision making cannot be stopped at this point. When there is a need to support automatic answer, decision makers search through their knowledge information that could support the automatic answer. If knowledge comes with a strong positive feeling (FOR), a decision maker accepts the answer with a high confidence that the answer is right. However, when knowledge arrives with weak FOR, there arises a need to initiate the process of data collection and analysis, so as to support the automatic answer. This is a situation when an automatic answer appears with high FOR, but its explanation comes with low FOR. A probable reason of low FOR occurrence at the moment of explaining automatic answer is that a decision maker doesn't have knowledge that will overlap implicit structures based on which automatic answer with high FOR has occurred.

This decision-making path is characterized by a high importance of an initial feeling in the process of decision making. The initial strong feeling causes decision maker to accept automatic answer as the right one. All the process of searching knowledge, collecting and analyzing data appears so as to support the automatic answer and to explain away possible contradictory information. This initial feeling highly reduces effort, because the process of data collection is focused only on one direction (to support automatic answer). The initial feeling appears to be problematic in situations when automatic answer occurs to be incorrect and the high feeling precludes decision maker from detecting information helpful in finding the correct answer.

Decision-making path 2: automatic answer occurrence with low FOR.
The second decision-making path describes a situation when an automatic answer occurs with low FOR. Low FOR informs that there are implicit structures that vaguely fit to the given decision situation. In the case of executive search consultants, it could mean that the candidate resembles to some extent successfully accepted candidates for similar positions. If there is no need or possibility to support automatic answer, decision makers accept the answer with a low confidence that the answer is right. The confidence is low, because low intensity of the initial feeling allows doubts occurrence. When there is a need to support automatic answer, decision makers search through their knowledge information that could support or deny the automatic answer. If support or denial of the automatic answer arrives with high FOR, decision makers accept it with a high confidence the decision is right. But, when knowledge arrives with weak FOR, there arises a need to initiate the process of data collection and analysis, so as to support or to deny the automatic answer.

This decision-making path is characterized by a lower role of the initial feeling in the process of decision making as in the previously explained path. This process reminds decision making defined as rational. A decision maker engages similar effort in the process of searching information that support and deny automatic answer. The drawback of this decision-making path is that it is much more effortful than the one previously described.

Decision-making path 3: no automatic answer occurrence.
The third decision making path describes a situation when no automatic answer occurs. Based on the concept of meta-reasoning, this situation means that there are no implicit structures that fit to the decision situation. In the case of executive search consultants, it means, that the candidate doesn't resemble previously accepted candidates. In such a situation, a decision maker needs to collect more information until the moment automatic answer occurs.

DISCUSSION

The presented study on expert executive search consultants' process of decision-making allowed to see how intuition and a corresponding Feeling of Rightness support consultants' process of reasoning and decision-making. The results of the study indicated that the initial intuitive appraisal highly influences final decision about accepting a candidate for a searched position. The initial intuition was found also to influence the way consultants were collecting and interpreting information. Results of the study indicated that in situations when consultants had a high positive feeling towards a candidate, they were focused more on collecting supporting than rejecting information. Moreover, they tended to explain away contradictory information even though they concerned fundamental requirements. Similarly, but in reverse it looked in a situation where there was a strong negative impression. In situations where consultants did not have any strong impression, their form of collecting information was focused neither on supporting nor on denying candidate's application. As the consultants have admitted that their initial feeling towards a candidate arises due to their extensive experience, it could be assumed that the analyzed feeling is the Feeling of Rightness (FOR) described by the theory of Meta-Reasoning.

The described study has an exploratory form. Results of the study gave a base to create experts' decision-making model that is compatible with the concept of meta-reasoning. However, in order to approve validity of the model, it should be verified on a bigger sample of experts and be tested in form of experiments.

Results of the study would be helpful to design a Decision Support System (DSS) for intuitive decision-making. The DSS could be applied in organizations where members need to make decisions in the absence of clear facts and intuition is usually the only form of finding solutions. The role of the system would be to measure probability of correctness of intuition and to detect misleading intuitions. In order to design such a system, presented results of the study should be verified on different groups of experts and in form of experiments. Afterwards, tools that measure probability of correctness of intuition should be designed. The first tool will estimate the level of task relevant experience of a decision-maker. As intuition is perceived to be based predominately on implicit knowledge (Underwood, 1996), the tasks should measure not only explicit but also implicit knowledge.
The second tool will estimate the level of intensity of Feeling of Rightness occurrence. Because of individual differences in the level of trusting one's intuition (Epstein et al., 1996), the system should verify automatically the intensity of the FOR. In order to design such a measurement, a scale of the intensity of FOR occurrence should be designed. Based on the scale, the DSS could use skin conductance sensor in order to verify to what extent the occurring intuition should be trusted. In order to implement DSS for intuitive decision-making effectively, members of organizations should acquire relevant knowledge about intuition and its role in decision-making.

ACKNOWLEDGMENTS

I would like to thank Dariusz Użycki PhD for helping me find expert executive search consultants for the study and for introducing me to the work of executive search consultants. I would like also to thank Gary Klein PhD and Gordon Pennycook PhD for providing valuable comments about the model presented in the paper.

REFERENCES

Ackerman, R., Thompson, V.A. (2017). Meta-Reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21(8).*

Britton, L.C., Wright, M., Ball, D. (2000). The use of co-ordination theory to improve service quality in executive search. *The Service Industries Journal, 20(4),* p. 85.

Clerkin, T.A., Lee, J. (2010). Executive search relationships-contacts between executives and search firm professionals: Scale development and validation. *Organization Management Journal, 7(3),* pp. 208-228.

Dane, E., Pratt M., (2007). Exploring Intuition and its role in managerial decision making. *Academy of Management.*

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, 7, pp. 28-38.

De Neys, W. (2017). Dual Process Theory 2.0. *A Psychology Press Book.*

Dreyfus, S.E. (2004). The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science Technology and Society.*

Epstein, S., Pacini, R., Denes-Raj, V. and Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, Vol. 71 No. 2*, pp. 390-405.

Evans, J.St.B. (2017). Dual Process Theory. Perspectives and problems. In: W. De Neys. Dual Process Theory 2.0. A *Psychology Press Book*, pp. 137-156.

Evans, J.St.B. (2011). *Dual-process theories of reasoning: Contemporary issues and developmental applications*, Developmental Review, 31, pp. 86-102.

Evans, J.St.B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 69, pp. 1-17.

Finlay, W., Coverdill, J. (1999). The search game: Organizational conflicts and the use of headhunters. *Sociological Quarterly, 40(1),* pp. 11-30.

Finucane, M.L., Alhakami, A., Slovic, P., Johnson, S.M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 12*, pp. 1-17.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19, pp. 25-42.

Greer, C.R., Youngblood, S.A., Gray, D.A. (1999). Human Resource Management Outsourcing: The Make or Buy Decision. *The Academy of Management Executive, 13,* pp. 85-96.

Guthridge, M., Komm, A.B., Lawson, E. (2008) Making talent a strategic priority. *McKinsey Quarterly,* (1), pp. 48-59.

Miles, A., Eugene Sadler-Smith, E. (2014).”With recruitment I always feel I need to listen to my gut”: the role of intuition in employee selection. *Personnel Review*, Vol. 43 Iss 4 pp. 606 – 627.

Newstead, S.E., Pollard, P., Evans, J.S. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition, 45*, pp. 257-284.

Rao, H., Drazin, R. (2002). Overcoming Resource Constraints on Product Innovation by Recruiting Talent from Rivals: A Study of the Mutual Fund Industry, 1986-94. *Academy of Management Journal, 45(3),* pp. 491-507.

Stanovich, K.E. (1999). Who is rational? Studies of individual differences in reasoning. *Lawrence Erlbaum Associates Publishers, Mahwah, NJ.*

Thompson, V.A., Turner, P.J.A., Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology 63 (3)*, pp. 107-140.

Underwood, G. (1996). Implicit Cognition. *Oxford Science Publication*.

Wong, B.L.W. (2004). Critical Decision Method Data Analysis. In D. Diaper & N. Stanton (Eds.), Task Analysis for Human- Computer Interaction, *Mahwah, NJ: Lawrence Erlbaum Associates*, pp. 327-346.

Development of a Practical Toolbox of Adaptation in Aviation and Air Traffic Control

Stathis Malakis[a], Tom Kontogiannis[b], Panagiotis Psaros[a]
[a]*Hellenic Civil Aviation Authority, Rhodes/ Diagoras International Airport, Rhodes Hellasa*
[b]*Cognitive Ergonomics & Industrial Safety Laboratory, Department of Production Engineering & Management, Technical University of Crete, Chania Hellas*

ABSTRACT

Adaptation can be defined as making appropriate responses to changing situations while adaptability enables people to recognize and adapt smoothly to unexpected situations. Commercial aircraft operations and Air Traffic Control (ATC) are complex and tightly coupled systems moving passengers and cargo with safety and regularity worldwide. Although they are highly scripted environments, adaptation remains one of the most-valued concepts. In this line of reasoning, we explored the concept of adaptation in the commercial aircraft operations and ATC through the lens of Resilience Engineering and Naturalistic Decision Making. We developed a model and a set of adaptation elements (i.e. a toolbox for adaptation) in order to explore the nature of adaptation as a source of resilience and develop a practical toolbox with wider applications in diverse organizations and challenging situations.

KEYWORDS

Adaptation; Aviation; Air Traffic Control; Resilience Engineering; Naturalistic Decision Making; Safety.

INTRODUCTION

Adaptation can be defined as making appropriate responses to changing situations (Hollnagel and Woods, 2005). Paradoxically, people at the sharp-end of organizations can recognize and intuitively understand cases of adaptation but there has been little guidance so far on how to learn, practice, evolve and transfer their experiences. A common perspective claims that practitioners constitute a general adaptive capacity for organizations compensating for shortfalls and problems at any level and that adaptation emerges as a result of their own training and accumulated operational practice.

Commercial aircraft operations and Air Traffic Control (ATC) units are complex and tightly coupled systems, moving passengers and cargo with safety and regularity worldwide. ATC is a work domain that relies on the cognitive functions of controllers and their collaboration with flight crews, airport operators, and other stakeholders to control the airspace, manage safety and adapt to the changing demands of everyday operations (Kontogiannis and Malakis, 2017). Nowadays, the tasks of flight crews in commercial aviation have seen a notable change from flying the aircraft by manual control, to increased system monitoring in the cockpit. Flight automation systems enable stability of operations and management of variations and disturbances, leaving flight crews mainly with the tasks of decision-making and system-monitoring. However, flight automation does not eradicate other sources of variation such as complex interactions with ATC, ground delays, airport congestion and ground handling issues. Therefore, a high level of variation is still present in everyday operations.

Although commercial aviation and ATC are highly scripted environments, adaptation remains one of the most-valued concepts. It is commonplace for airlines and ATC to assert that they encourage their sharp-end people (i.e. flight crews and air traffic controllers) to be adaptive. They place a high value on having adaptive teams, as opposed to rigid teams, because the uncertainty of everyday operations requires people to be able to adjust to changing conditions. Yet their approach relies mostly on over-utilization of known solutions such as, adding more procedures, embracing standardization and promoting a multiplication of rules. There is a discernible tension between 'rule compliance' and 'adaptation' that has been reported by many practitioners in their daily work. For instance, McDonald (2006) summarized the results of a series of European projects about aircraft maintenance and concluded that front-line operators would find 'better, quicker and even safer ways of doing tasks' than following procedures routinely.

At the operational level, commercial aviation and ATC share many common work demands that are typical of safety-critical systems. At a meso-level, operations are characterized by multiple interdependent control centres, each with partial authority and autonomy but with different goals in mind. At a higher-level, aviation and ATC have been influenced by two competing goals: (i) Higher throughput and (ii) Lower accident rates. Although a lot of emphasis has been given to higher throughput, safety remains the fundamental mandate of aviation and ATC operations (ICAO, 2006; 2018).

According to Morel, Amalberti and Chauvin (2008), the safety of complex systems comprises two aspects, that is, the part of safety achieved through rule, prohibitions, limitations, legal requirements (referred to as Controlled

Safety) and the part of safety supported by the adaptive intelligence of the operators and professionals (Managed Safety). Systems that achieve ultra-safe safety status- through long, costly and dedicated efforts - do not rely on the adaptive expertise of their practitioners; instead, they require them to be identical and interchangeable individuals within their respective roles, equipped with standardized competencies and assisted by organizational structures and automation safety nets. The adaptive capacity of their people is constrained because, firstly they are rarely exposed to exceptional situations and secondly relevant guidance is minimal.

It is widely accepted that training does not adequately prepare flight crews to cope with demanding scenarios with ambiguous and potentially conflicting information. Burian, Barshi, and Dismukes (2005) has shown that refresher training in aviation provided limited opportunities to flight crews to practice emergency and abnormal situation procedures in the context of real-world situations. The same training philosophy applies to the ATC where simulator time is dedicated to regulatory compliant exercises. Malakis and Kontogiannis (2012) analyzed a set of real incidents regarding 'separation minima infringements' and simulated events encountered in refresher courses. A comparison of the simulator training sessions and the real events indicated that controllers excelled in managing textbook emergencies in simulator training but different performance patterns emerged in the operations room. When confronted with real traffic, even expert controllers made errors in recognizing problems, anticipating threats, and planning the traffic especially when they were working under heavy workload. It is commonplace for front-end operators to get minimal training and guidance to work outside this procedural framework. In this sense, aircraft operations and ATC gravitate towards Controlled Safety and a culture of persistent minimization of Managed Safety.

In general, it has been very difficult to preserve both the expertise of practitioners in abnormal situations and the benefit of achieving maximum safety by procedural means (Amalberti, 2013). In aviation and ATC, however, several examples exist of extraordinary episodes of adaptation. For instance, the "Miracle of Hudson river" (i.e. US Airways Flight 1549 successful ditching at Hudson River in 2009 without any casualties) and the management of air traffic during ATC Zero on September 11 2001 (when controllers managed safely the grounding of thousands of aircraft in the US skies at a short notice). These two celebrated episodes of adaptation have been attributed to the airmanship of flight crews and professionalism of air traffic controllers. In this line of reasoning, we explored the concept of adaptation in the commercial aircraft operations and the ATC through the lens of Resilience Engineering and Naturalistic Decision Making. In the following sections, we attempt to make sense of and provide a practical toolbox for adaptation. Our aim was to explore the nature of adaptation as a source of safety in ATC and commercial aircraft operations.

LITERATURE REVIEW

The Human factors literature is replete with examples of practitioners adapting to system changes and unforeseen events in complex systems. Adaptations may be triggered by design flaws or modern technology failures with practitioners trying to come up with alternative ways to the intended use of technology (i.e., workarounds or kludges) or finding new ways to compensate for design flaws or component failures (Koopman and Hoffman, 2002; Kontogiannis and Malakis, 2013). In cybernetic terms, adaptations can be seen as efforts to increase the 'variety of practitioner's to match the demands of 'process variety' (e.g., fluctuations, changes, disturbances). High 'process variety' makes performance outcomes difficult to predict, as alternative plans may lead to the same goal and the same actions do not always lead to the same outcome (Brehmer, 1992).

Other studies of complexity-coping management have focused on cognitive strategies used by teams to recognize, make sense, or respond to process variations (Kontogiannis, 2010; Kontogiannis and Malakis, 2012; Furniss, Back, Blandford, Hildebrandt and Broberg, 2011). In a study of air traffic controllers in coping with complexity, Kontogiannis and Malakis (2012) developed a set of behaviour markers of adaptation in monitoring and anticipation, re-planning, re-structuring tasks and re-organizing communication. Later, this list of behaviour markers of adaptation was expanded to a number of team processes such as, common operating picture, transfer of control and management of margin of manoeuvre.

Another way to describe adaptations has been through the principle of 'approximate adjustments' (Kontogiannis and Malakis, 2013). As the work conditions rarely match the ones specified in procedures, practitioners are constantly making adjustments to meet the demands of the situation at hand. Because the resources are finite and the window of opportunity is short, adjustments are always 'approximate' rather than 'precise'. These 'approximate adjustments' include, informal work practices that extend job prescriptions and 'leverage points' whereby practitioners exploit opportunities at work to optimize their performance (Malakis and Kontogiannis, 2020).

Resilience engineering provides yet another view of adaptation compared to the traditional human factors literature where adaptations or workarounds are seen as indicators of a poor fit between technology-procedures and actual work conditions (Koopman & Hoffman, 2002). The resilience perspective looks into the capabilities of the system to cope with increasing demands and compensate by adapting its performance. Hence, the focus shifts from individuals to teams and organizations who have to cope with multiple goals, organizational pressures, and complexity. Resilience engineering is about understanding and anticipating what sustains and what erodes adaptive capacity (Patterson, Woods, Cook and Render, 2006). To be resilient, a system needs to be able to anticipate what may happen, monitor what is going on, respond effectively when something happens, and learn from past experiences (Hollnagel, 2018). In this sense, adaptation is considered as a 'composite' rather than a basic 'potential' of resilience. Hollnagel argues that adaptation is a combination of the potentials to learn, to respond and possibly to monitor (Hollnagel, 2018; p49); hence adaptation is not figured in any visible manner.

Woods (2018) presented a detailed introduction of the 'theory of graceful extensibility', which expresses the fundamental characteristics of adaptability. 'Graceful extensibility' refers to the capability of people to stretch at or near the system boundaries and continue to deliver system outcomes, when surprises occur. A central premise of the theory of graceful extensibility is that resilience will emerge when individual 'units of adaptive behavior' exhaust their own adaptive capacity and obtain additional capacity from other units in their network.

Adaptations have been seen as tailoring technology to the needs of the problem, as cognitive strategies for adapting human performance to complex systems, as approximate adjustments and leverage points, as matching human variety to the process variety, as graceful extensibility of human behaviour and finally, as a system capability to compensate to increasing system demands. These perspectives are not necessarily divergent and they may provide a good basis for developing a new perspective into the individual and team processes of adaptation.

Klein and Pierce (2001) offered some challenging assertions about the nature of team adaptation, and the factors that contribute to effective performance. They specified several input factors that affect team adaptation such as, having the big picture so that local goals do not overtake team objectives, being able to control more degrees of freedom, or to support self-organizing efforts (decentralization), having team competence at everyday routines, being competent at solving problems as a team and so on. Klein and Pierce (2001) also specified several 'emergent states' (e.g., common ground and big picture) and 'adaptation processes' such as, having a mindset for adaptation, putting in action workarounds, figuring out when common ground has been lost, figuring out how to decompose a task in a different way, visualizing a plan in progress, estimating how long it will take to put changes into action, performing a trade-off between the advantage gained by improving a plan versus the disruption caused and so on. Finally, they specified a problem-solving mechanism operating on the emergent states and adaptation processes involving the stages of problem detection, re-framing and re-planning or revision of plans.

Woods and Branlat (2010) have proposed a framework of polycentric control to manage the adaptive capacity of a system. They seemed to focus on general system adaptation processes that relate to keeping pace with the escalation of the problem, working at cross-purposes and updating or revising interpretations and plans of action as the situations takes a different turn. In particular, Woods and Branlat (2011) looked into several breakdowns of adaptive performance that could be assigned into three categories, that is: (i) decompensation failures, (ii) problems at working at cross-purposes and (iii) getting stuck in outdated behaviours.

Two particular studies that looked into the adaptation processes in ATC regard the analytic framework of Rankin, Lundberg, Woltjer, Rollenhagen, & Hollande (2014) and the influential factors of adaptation of Foster et al. (2019). In particular, Rankin et al. (2014) presented a framework to analyse adaptive behaviours and a variety space diagram (i.e., a space of possible interpretations and actions) as a means to illustrate how system variability, disturbances, and constraints can affect work performance. Foster, Plant and Stanton (2019) presented a nine-factor model of adaptation in complex socio-technical systems with a case study in the ATC. In the validation phase, they applied the model to explain the response of UK oceanic controllers to the closure of US airspace on the aftermath of the terrorist attacks in the US on the 11th September 2001.

By reviewing earlier research, we created a synthesis of the findings that would be appropriate in the context of commercial aircraft operations and ATC.

METHOD

To make sense of the term 'adaptation' we analysed 'adaptive failures and successes' in order to get a rich picture and a deeper understanding of the concept. The three sources of data we used are the following:

- An analysis of 109 cases of 'separation minima infringements' worldwide (1996-2016). These represent cases of adaptive failures and focus mostly on ATC operations although they provide valuable insights for the aviation domain.
- An analysis of five aviation accident investigation reports that represent well-known cases of successful and unsuccessful adaptation, that is: Air France Flight AFR447 (BEA, 2012); Air Transat Flight 236 (Air Accident Investigation, 2004); US Airways Flight 1549 (NTSB, 2009); Air Canada Flight 143 (Air Accident Investigation, 1985) Quantas Flight QF32 (ATSB, 2013; de Crespigny, 2012).
- An analysis of a large number of selected ATC simulator, real time scenarios and aviation accident investigations made by the two authors on the ATC and the aviation sectors over the past 10 years (Malakis et al. 2010a ; 2010b; Kontogiannis, and Malakis, 2012; Malakis and Kontogiannis, 2020; Malakis, Psaros, Kontogiannis, and Malaki, 2020).

RESULTS

Adaptation of performance is a broad concept that refers to many cognitive and team processes. In this sense, many different performance models could be used to examine how people adapt their behaviours and strategies prescribed in the particular models. We used a command and control framework (i.e., Taskwork & Teamwork in

Effective and Adaptive Management, T^2EAM) that has been developed in the context of ATC (Kontogiannis and Malakis, 2017). In this sense, the T^2EAM model (Figure 1) provides a basis for studying a variety of adaptation elements that have been derived from the data analysis.

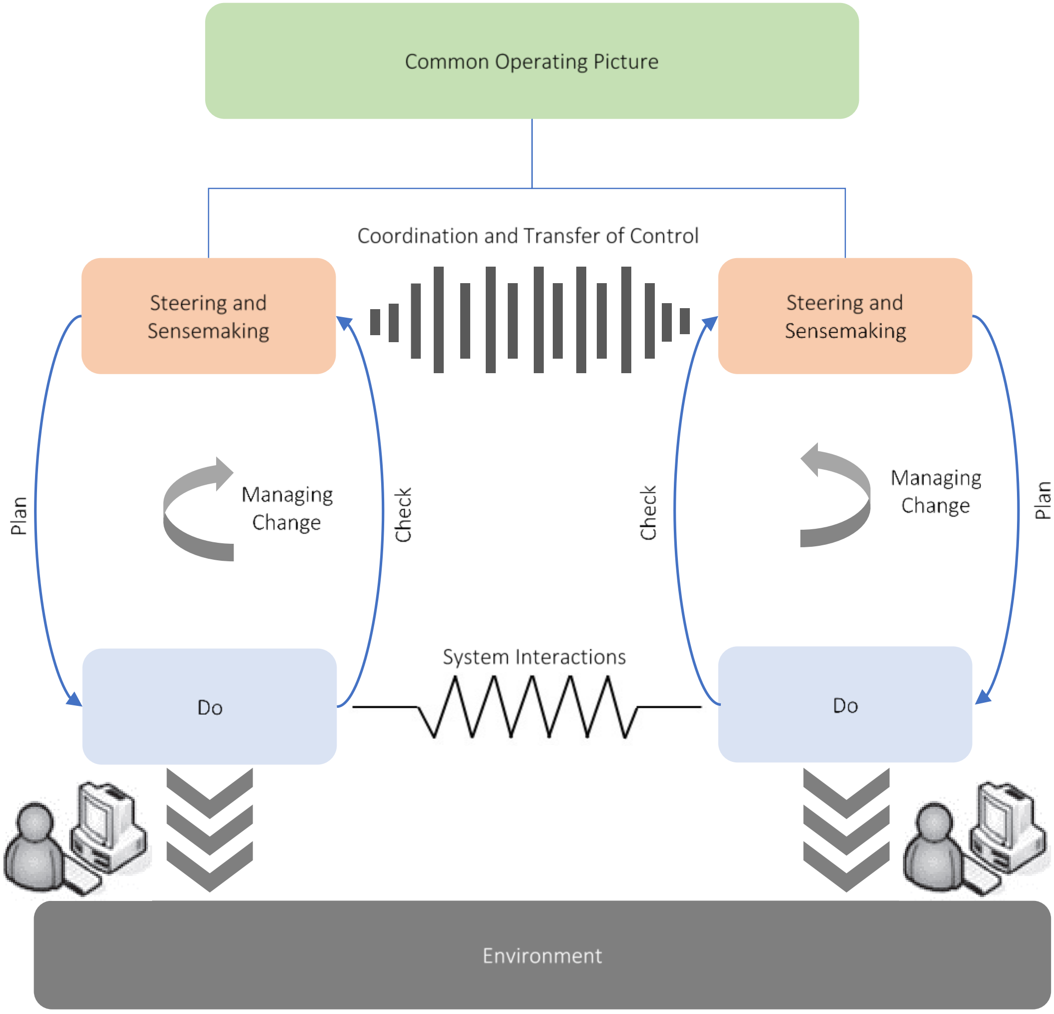


Figure 1. A conceptual model of the proposed toolbox of adaptation

In particular, this model examines six cognitive and team processes as follows:

- Steering or goal-setting.
- Sensemaking and Mental Models.
- Common Operating Picture or shared mental model.
- Coordination and transfer of control.
- Managing changes.
- Planning-Doing-Checking cycle.

Based on the literature review and the analysis of data we developed a set of adaptation elements (i.e. a toolbox for adaptation) that accounts for all cases of successful and poor adaptation recorded in the set of data.
The toolbox for adaptation is shown in Table 1.

Cognitive and Team Processes	Short Description	Elements
Steering	An orientation of performance towards specific goals that can be modified later in the course of events	• Willingness to make the necessary trade –offs. • Spotting and exploiting leverage points. • Making enlightened experimentations.
Sensemaking and Mental Models	A process of interpretation of the situation to make sense of the problem and understand the factors that contributed to its occurrence. Sensemaking relies on the mental models of practitioners that are developed and refined over time	• Refining mental models. • Discerning temporal and long-term trends. • Unwillingness to 'explain away' conflicting data • Reconfiguring an operationally viable system within a degraded system.
Common Operating picture	A COP refers to a common perception of threats, available resources and forces, opportunities for action and assessments of work.	• Building a common operating picture. • Communicating intent.
Coordination and Transfer of Control	Another team process that has been included in order to examine the coordination of multiple loops or units, the handover of the situation between shifts and the interaction of people and automation	• Reconfiguring task and priorities with minimum coordination cost. • Cross checking and monitoring performance of other team members. • Maintaining an uninterrupted information flow.
Managing Change	A process for addressing changes in the system or the environment that may have an impact on performance	• Maintaining Margin of Maneuver. • Synchronizing speed of responses with escalation rate.
Planning – Doing - Checking	A basic performance cycle that involves devising a plan to implement a certain goal, executing the plan and monitoring its effects.	• Anticipatory thinking. • Bending operational procedures and emergency checklists. • Working around a problem. • Choosing when to change from linear to modular plans. • Establishing task triggers and error barriers. • Coping with weak signals and masking effects.

Table 1. A toolbox for adaptation

Steering

Steering broadly refers to the choices and goals that people set in order to achieve their mission. To be successful in steering, people need to demonstrate willingness to make the necessary trade-offs and show an ability to spot and exploit leverage points. Fundamental to steering is the presence of a large 'database of action repertoires' in the form of work practices honed through 'enlightened' experimentation (Thomke, 2001). The following elements of adaptation could be considered:

- Willingness to make the necessary trade -offs. Trade-offs between goals may be apparent but there may be a reluctance to resolve them.
- Spotting and exploiting leverage points. Experts notice 'interesting events' (that is a 'leverage point') that provide opportunities for building efficient courses of action that make a significant impact on the operation of the system.
- Making 'enlightened' experimentation: Work practices are usually developed through experimentation and often may lead to performance improvements.

Sensemaking and mental models

Sensemaking relies on having an adequate 'mental model' of potential hazards, causes, available resources and risk control strategies. In ATC and aviation, a common problem is 'failure to revise' a mental model or mindset, as new evidence becomes available. The initial situation assessment may seem appropriate, given the availability of information, but controllers and flight crew may fail to revise their mindsets. Many patterns of breakdown relate to problems in managing complexity and flawed mental models (Grote, 2009). The following elements of adaptation could be considered:

- Refining mental models: Controllers and pilots develop in depth operational knowledge over a period of years.
- Discerning between temporal and long-term situational trends. This adaptation element is most prominent in dealing with situations of progressive weather deterioration or systems failures.
- Unwillingness to 'explain away' conflicting evidence: Misleading cues combined with unusual indications may create an unforgiving environment that inhibits detection of erroneous hypotheses.
- Re-configuring a basic operational system from a largely degraded system: In the case of Quantas flight 32, captain de Crespingy managed to "build" a simple operational plane from the larger ailing Airbus 380 (de Crespigny, 2012).

Common Operating Picture

The COP is an important frame for making sense of a problem in a collective manner, along the ideas put forward by Karl Weick (1995). As reported by Luokkala, Nikander, Korpi, Virrantaus and Torkki (2017), the COP should incorporate information that enables situational information to be produced, visualized and presented in such a way that all information is available to all actors involved in the crisis in real time. In transforming data into a COP and make sense of the situation, pilots and controllers remain adaptable through the following elements:

- Building and disseminating a common operating picture: In order to achieve a continuum of synchronized action, teams must meet the emerging need for maintaining and disseminating where appropriate a COP.
- Communicating intent: This element refers to the extent that controllers and pilots communicate ambiguously and clarify their intentions.

Coordination and transfer of control

Coordination and transfer of control is a complex process that involves team synchronization, handover of the situation between shifts and interaction with colleagues and automation. The very nature of teamwork and the allocation of responsibilities can generate many dependencies that require orchestrated action in order to converge them all toward the same goal.

- Re-configuring tasks and priorities with minimum coordination cost: This element refers to the ability of controllers and pilots to adjust the delegation of tasks and priorities.
- Efficiently crosschecking and monitoring performance of others: Controllers and pilots employ efficient monitoring strategies that enable them to catch errors and correct and/or provide feedback without hindering the work of their teammates.
- Maintaining an uninterrupted and ungarbled information flow: Keeping the size and duration of communication to the practical minimum is an important feature of adaptive teams.

Managing changes

Adaptive teams monitor changes in the system or the environment and try to match their capabilities to them or even reserve their capabilities for anticipated events. The implication is that practitioners should retain some residual capacity for managing a number of secondary activities that have to do with correcting side-effects and coping with task interruptions, including an assessment of one's own capabilities and "margins for maneuver."

- Maintaining 'Margin of Maneuver': Adaptive teams proactively monitor changes in the Margin of Maneuver over time and regulate it in anticipation of potential challenges to assure that they possess the required resources.
- Synchronizing speed of responses with escalation rate: Practitioner' activities are mainly driven by their recognition, planning, and coordination demands in order to match the tempo and intensity of work. The transition between normal and high tempo operations can be rapid and practitioners should synchronize their speed of responses with the escalation rate.

Planning – Doing – Checking cycle

The basic planning-Doing-Checking cycle underlies all types of performances where people devise a sequence of actions and use their resources to achieve a goal within certain constrains, executing the plan in a timely fashion and checking over or evaluating their progress of work with regard to the goals set earlier.

- Anticipatory thinking: According to Klein, Snowden and Pin (2010), this refers to the ability to prepare in time for problems and opportunities. It is different from prediction because practitioners are preparing themselves for future events, not simply predicting what might happen.
- Bending operational procedures and emergency checklists: In the case of Quanta's flight 32, the flight crew intentionally skipped some steps of the emergency checklists. Faced with a similar situation they thought of the Air Transat flight 236 in which that crew followed the checklist procedures and transferred all the fuel from the good wing to the leaking one and skipped that element of checklist (de Crespigny, 2012).
- Working around a problem to find a temporal solution: In complex engineered systems, work issues cannot be specified in procedures and remain to be resolved by people at the sharp end (Woods and Cooke, 2002). Work practices capture the cognitive flexibility that experts exhibit in work contexts characterized by complexity and uncertainty (Feltovich et al. 1997). Many times, practitioners need to utilize established workarounds to a problem in order to find temporal solutions and keep the situation under control.

- Changing from linear to modular plans: Klein and Pierce (2001) showed that detailed and highly coupled plans are more difficult to adapt in-the-moment relative to modular plans. Normally, the planning may be linear before the need for adaptation is triggered. When the adaptation process starts, however, there is an urgent need to revert to modular planning.
- Establishing reminders, task triggers and error barriers: Reminders and task triggers can help people detect omissions, particularly in cases where tasks are independent from each other. An alternative way to combat this natural tendency is to create a barrier so that errors are stopped before having an adverse consequence.
- Coping with masking effects: Numerous technical systems have long response times and masking effects that make the problem difficult to detect or interpret (Kontogiannis, 1999).

CONLCUSION

This study has sought to develop a model of team adaptation and a practical toolbox for looking into the adaptation elements from studies in ATC and commercial aviation. We developed a model of team adaptation that includes both cognitive and team processes such as: steering, sensemaking, common operating picture, coordination and transfer of control, managing changes and the planning-doing-checking cycle. Our aim has been to explore the nature of adaptation as a source of resilience in the ATC and aviation domains and develop a practical toolbox applied to a variety of organizations and range of work situations. In this sense, our study appears to make four contributions: (i) insights into how adaptations made by operating teams can support resilience in case of critical events (ii) a toolbox of adaptation mechanisms that could be used in error-management training programs (iii) insights into the improvement of the quality of information in investigating challenging events.

The model and the toolbox for adaptation aim to analyse past events and future challenging scenarios and hence, support both retrospective and prospective safety management efforts. It is clear that new analytical methods, that may not have been traditionally applied in the human factors and safety domain, are needed to make sense of this quantitative data with the leverage of subject matter experts. Initial ideas in this respect have been explored for air traffic management by Malakis et al. (2020). Further studies are required to evaluate the proposed model and toolbox of adaptation as a means to analyse adaptive strategies and as a guidance material to integrate with existing safety management activities. This work provides an initial framework for researchers and practitioners to identify critical adaptation enablers that may not surface through traditional human factors and safety analysis methods.

REFERENCES

Air Accident Investigation, (1985). *Final Report of the Board of Inquiry lnvestigating the circumstances of an accident involving the Air Canada Boeing 767 aircraft C-GAUN that effected an emergency landing at Gimli, Manitoba on the 23rd day of July, 1983.* Canadian Government Publishing Centre Supply and Services. Canada.

Air Accident Investigation, (2004). *All engines out landing due to fuel exhaustion; Air Transat Airbus A330-243 C-GITS, Lajes, Azores, Portugal, 24th August 2001.* Portuguese Aviation Accidents Prevention and Investigation Department. Accident investigation final report.

Amalberti, R. (2013). *Navigating Safety: Necessary Compromises and Tradeoffs.* New York: Springer.

ATSB. (2013). *In-flight uncontained engine failure overhead Batam Island, Indonesia 4 November 2010, VH-OQA Airbus A380-842. Aviation Occurrence Investigation – AO-2010-089 Final.* Australian Transport Safety Bureau. 27 June 2013.

BEA. (2012). *Final Report On the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro - Paris.* Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile. Aéroport du Bourget. France.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*. 81(3): 211–241.

Burian, B. K., Barshi, I., & Dismukes, K. (2005). *The Challenge of Aviation Emergency and Abnormal Situations.* Technical Report No. NASA/ TM—2005-213462. Moffett Field, CA: NASA Ames Research Center.

De Crespigny, R. (2012). QF32. Pan Macmillan Australia, Australia.

Feltovich, P. J., Spiro, R. J., and Coulson, R. L. (1997). Issues of flexible flexibility in contexts characterized by complexity and chance. In P. J. Feltovich, K. M. Ford, and R. R. Hoffman (Eds.), *Expertise in Context: Human and Machine.* Cambridge, MA: MIT Press, 125–146.

Foster, C. J. Plant, K. L. and Stanton, N. A. (2019). Adaptation as a source of safety in complex socio-technical systems: A literature review and model development. *Safety Science,* 118: 617–631.

Furniss, D. Back, J. Blandford, A. Hildebrandt, M. & Broberg, H. (2011). A resilience markers framework for small teams. *Reliability Engineering & System Safety*, 96, 2–10.

Grote, G. (2009). *Management of Uncertainty: Theory and Application in the Design of Systems and Organizations.* Springer, Berlin

Hollnagel, E. & Woods, D. D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering.* Boca Ranton: CRC Press, Taylor & Francis Group.

Hollnagel, E. (2018). *Safety-II in Practice: Developing the Resilience Potential.* Routledge.

ICAO. (2006). *Convention on International Civil Aviation Organization.* Document 7300/9. Ninth Edition. Montreal: International Civil Aviation Organization.

ICAO. (2018). *Safety management manual (SMM).* Document 9859, AN/474, 4th edition. International Civil Aviation Organization, Montreal.

Klein, G. & Pierce, L. (2001). *Adaptive Teams.* 6th International Command and Control Research and Technology Symposium. US Naval Academy, Annapolis MD.

Klein, G. Snowden, D. & Pin, C. L. (2010). Anticipatory thinking. In K. Mosier & U. Fischer (Eds.), *Informed knowledge: Expert performance in complex situations.* London, UK: Psychology Press.

Kontogiannis, T. & S. Malakis. (2013). Remaining safe by working at the edge of compliance and adaptation: Reflective practices in aviation and air traffic control. Theoretical Issues in Ergonomics Science, 14 (6): 565–591.

Kontogiannis, T. (1999). User strategies in recovering from errors in man machine systems. *Safety Science,* 32:49–68.

Kontogiannis, T. (2010). Adapting plans in progress in distributed supervisory work: Aspects of complexity, coupling, and control. *Cognition Technology & Work,* 12: 103–118.

Kontogiannis, T., & S. Malakis. (2012). Strategies in coping with complexity: Development of a behavioural marker system for air traffic controllers. *Safety Science,* 57: 27–34.

Kontogiannis. T. & Malakis, S. (2017). *Cognitive Engineering and Safety Organization in Air Traffic Management.* Taylor & Francis Group, Boca Raton.

Koopman, P. & Hoffman, R.R. (2002). Work-arounds, make-work, and kludges. IEEE Intelligent Systems–Human Centered Computing, 18(6):70–75.

Luokkala, P. Nikander, J. Korpi, J. Virrantaus, K. & Torkki, P. (2017). Developing a concept of a context-aware common operational picture. *Safety Science*, 93: 277–295.

Malakis, S. and Kontogiannis, T. (2012). Refresher training for air traffic controllers: Is it adequate to meet the challenges of emergencies and abnormal situations? *International Journal of Aviation Psychology* 22(1): 59–77.

Malakis, S. and Kontogiannis, T. (2020): Leverage points: insights from a field study in the air traffic control system. *Theoretical Issues in Ergonomics Science*, DOI: 10.1080/1463922X.2020.1784307.

Malakis, S. Kontogiannis, T. & Kirwan, B. (2010a). Managing emergencies and abnormal situations in air traffic control (part I): taskwork strategies. *Applied Ergonomics*, 41: 620–627.

Malakis, S. Kontogiannis, T. & Kirwan, B. (2010b). Managing emergencies and abnormal situations in air traffic control (part II): teamwork strategies. *Applied Ergonomics,* 41: 628–635.

Malakis, S., Psaros, P., Kontogiannis, T. and Malaki, C. (2020). Classification of air traffic control scenarios using decision trees: insights from a field study in terminal approach radar environment. *Cognition, Technology & Work 22*, (1): 159-179.

McDonald, N. (2006). Organizational resilience and industrial risk. In: E. Hollnagel, D.D. Woods, and N. Leveson, eds. *Resilience Engineering: Concepts and Precepts.* (pp.205–221) Aldershot: Ashgate.

Morel, G. Amalberti, R., & Chauvin, C. (2008). Articulating the differences between safety and resilience: the decision-making of professional sea fishing skippers. *Human Factors* 50(1):1–16.

NTSB. (2009). *Aircraft Accident Report, Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549, Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009.* National Transportation Safety Board. May 4, 2010. NTSB/AAR-10/03.

Patterson, E. S. Woods, D. D. Cook, R. I. & Render, M. L. (2006). Collaborative cross-checking to enhance resilience. *Cognition, Technology & Work*, 9: 155–162.

Rankin, A. Lundberg, J. Woltjer, R. Rollenhagen, C. & Hollande, E. (2014). Resilience in Everyday Operations: A Framework for Analyzing Adaptations in High-Risk Work. *Journal of Cognitive Engineering and Decision Making*; 8:78-97.

Thomke, S. (2001). Enlightened experimentation. The new imperative for innovation. *Harvard Business Review*, 79(2): 66–75.

Weick, K.E. (1995). *Sensemaking in Organizations*. Sage, Thousand Oaks, CA.

Woods, D. D. & Branlat, M. (2010). Hollnagel's test: being 'in control' of highly interdependent multi-layered networked systems. *Cognition, Technology & Work*, 12(2): 95-101.

Woods, D. D. & Branlat, M. (2011). Basic Patterns in How Adaptive Systems Fail. In E. Hollnagel, J. Pariès, D. D. Woods, & J. Wreathall (Eds.), *Resilience Engineering in Practice.* Farnham, UK: Ashgate.

Woods, D. D. and Cook, R. I. (2002). Nine steps to move forward from error. *Cognition, Technology & Work,* 4: 137–144.

Woods, D.D. (2018). The theory of graceful extensibility: basic rules that govern adaptive systems. *Environment Systems and Decisions,* 38:433–457.

Joint Meeting

15th Conference on Naturalistic Decision Making

9th Symposium on Resilience Engineering

Bouncing forward from global crises and challenges

21-24th June 2021

Toulouse, France
Columbus, Ohio
and around the world online

NDM, REA, FONCSI, ICSI

Resilience Engineering Association

Naturalistic Decision Making

ICSI
Institut pour une culture de sécurité industrielle

FONCSI
Fondation pour une culture de sécurité industrielle

ENAC
Ecole Nationale de l'Aviation Civile