# MACROCOGNITION: THE SCIENCE AND ENGINEERING OF SOCIOTECHNICAL WORK SYSTEMS

EDITED BY : Paul Ward, Robert R. Hoffman, Gareth E. Conway,
Jan Maarten Schraagen, David Peebles, Robert J. B. Hutton
and Erich J. Petushek

**frontiers** Research Topics

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

# MACROCOGNITION: THE SCIENCE AND ENGINEERING OF SOCIOTECHNICAL WORK SYSTEMS

Topic Editors:
**Paul Ward,** University of Huddersfield, United Kingdom
**Robert R. Hoffman,** Florida Institute for Human and Machine Cognition, United States
**Gareth E. Conway,** Defence Science and Technology Laboratory (Dstl), United Kingdom
**Jan Maarten Schraagen,** TNO Netherlands Organisation for Applied Scientific Research, Netherlands
**David Peebles,** University of Huddersfield, United Kingdom
**Robert J. B. Hutton,** Trimetis Ltd., United Kingdom
**Erich J. Petushek,** Michigan State University, United States

# Table of Contents

# Editorial: Macrocognition: The Science and Engineering of Sociotechnical Work Systems

Paul Ward[1]*, Robert R. Hoffman[2], Gareth E. Conway[3], Jan Maarten Schraagen[4], David Peebles[1], Robert J. B. Hutton[5] and Erich J. Petushek[6]

[1] The Applied Cognition & Cognitive Engineering Research Group, University of Huddersfield, Huddersfield, UK, [2] Florida Institute for Human and Machine Cognition, Ocala, FL, USA, [3] Defence Science and Technology Laboratory (Dstl), Porton Down, UK, [4] TNO Netherlands Organisation for Applied Scientific Research, Soesterberg, Netherlands, [5] Trimetis Ltd., Bristol, UK, [6] College of Human Medicine, Michigan State University, East Lansing, MI, USA

**Editorial on the Research Topic**

**Macrocognition: The Science and Engineering of Sociotechnical Work Systems**

The increasing complexity of work systems and changes in the nature of workplace technology over the past century have resulted in a substantial shift in the nature of work activities, from those predominated by physical labor toward more cognitively oriented work. Modern work systems have many characteristics that make them cognitively complex: They can be highly interactive; comprised of multiple agents and artifacts; information may be limited, contested, or distributed across space and time; problems can be unexpected and emergent; task goals are frequently ill-defined, conflicting, and dynamic; planning may only be possible at general levels of abstraction or require adaptive solutions; a considerable degree of proficiency or expertise is required; the stakes are often high; and problems usually involve uncertainty, time-constraints, and stress. To complicate matters further, cognition in complex work settings is typically constrained by broader professional, organizational, and institutional practices and policies, which themselves can be a moving target as work systems and organizations adapt to a constantly-changing landscape. These features of cognitive work present significant challenges to scientific methodology and theory, and to subsequent design of reliable work methods and the technologies that shape them.

Historically, philosophers and scientists have used divergent methods to understand the mental activities experienced during cognitive work at multiple levels of analysis. Some have examined cognition at an associative, contextual, functional, or holistic level, relying on naturalistic methods to understand the higher mental processes as they work in harmony during goal-directed behavior. Others have embraced experimental and computational methods and favored internal control over external validity, often reducing cognition to a psychology of fundamental acts, such as short-term memory access and action selection at the millisecond level.

More recently, Macrocognition has evolved as a complementary paradigm, focused on how cognition adapts to complexity, particularly in work settings (Klein et al., 2003). Macrocognitive researchers have studied the cognitive functions and processes associated with skilled, adaptive, collaborative, and resilient cognitive work in the context of the aforementioned complexities of sociotechnical work systems. Typically, this research has been carried out using cognitive task analytic techniques that draw on both naturalistic and experimental methods (e.g., Crandall et al., 2006). The primary goals of research in Macrocognition are to better understand cognitive adaptations to complexity, to increase our theoretical understanding of the organism–environment relations by studying the mapping between cognitive work and real-world demands, to better understand work-as-done rather than work-as-prescribed, work-as-imagined, or work-as-disclosed, and to promote use-inspired research capable of improving system performance and informing theory development (see for instance Schraagen et al., 2008).

The aims of this Research Topic are to showcase some of the exciting research on Macrocognition being conducted by cognitive scientists, cognitive ergonomists, and cognitive systems engineers, and to demonstrate the broad reach of this relatively new discipline. The opening paper, co-authored by one of the pioneers of Naturalistic Decision Making and Macrocognition, Klein and Wright, describes the evolution of this research and identifies some of the key drivers of the origin of Macrocognition. The paper highlights how this discipline has shaped our thinking about core cognitive processes, and our capabilities for developing training, decision support systems, and system design in complex and uncertain environments.

Four papers examine Macrocognition in traditional and non-traditional yet complex work domains. They present research at different levels of analysis using methods ranging from naturalistic techniques and interviews to simulations and experiments. Baber and McMaster demonstrate how UK police forces gather, frame, and share information as a means to coordinate incident response, and manage the associated uncertainties, risk, and resources. Collins et al. examine sports coaches' use of decision-making strategies. Their findings indicate that deliberation is often used as an immediate check on initial intuitions, which are heavily influenced by prior planning and experience level. Brouwers et al. use a novel, simulated rail control task to examine cue utilization. Their data suggest that individuals with greater cue utilization were more effective at routing trains while managing additional sources of cognitive load. Porat et al. report a series of studies that evaluate how many unmanned automata a single operator can supervise and control. They show that experienced operators were able to supervise around 15 systems with a moderate level of automation but can only control up to three effectively. Moreover, teams of operators generally performed better than individuals working alone.

Two papers investigate Macrocognition in team settings and organizational networks. Buchler et al. investigated the assumption that greater information sharing improves situation awareness and organizational effectiveness. Their data suggest that sending many messages can actually decrease the likelihood of attaining shared situation awareness. The similarity between team members in terms of their functions and initial situation awareness levels likely impacted these results, highlighting important issues for networked organizations. Fiore and Wiltshire synthesize a broad set of perspectives on how team cognition occurs in complex collaborative contexts, as well as

the artifacts and technology that support team performance. They provide diagnostic guidelines on studying the relationship between artifacts and team cognition and present implications for how to conceptualize team-supporting technology.

Three papers investigate the role of Macrocognition in design. Fadde presents a framework for translating macrocognitive research into the design of instruction to take place in the workplace. He presents a case study that applies macrocognitive training to baseball and highlights the challenges of embedding such training in the work setting. Goode et al. examine how the macrocognitive approach can inform system design, specifically how incident data can be translated into prevention strategies that address the systemic causes of accidents. They argue that the design process needs to be refined to focus design on monitoring and feedback mechanisms that support high-level decisions. Naikar and Elix suggest that to create work systems that are capable of adapting to complexity, all system elements need to be integrated into the design in a way that supports workers' ability to adapt their behavior and the environmental structure in order to handle novelty as well as familiarity. They present an integrated design approach aimed at facilitating system performance through adaptation.

The final paper, by Laurent and Bianchi, offers a critical view of Macrocognition and asks whether it should be distinguished from other forms of cognition. They echo earlier comments by Klein et al. (2003) that Micro- and Macrocognition present research at different levels and scales of analysis. They argue for the development of a multiscale model of cognition, in which context and cognition interact at multiple levels.

These articles demonstrate the diversity of perspectives and methods employed in research on Macrocognition, as well as the pragmatic focus of this research toward leveraging our understanding of how cognition adapts to complexity. We are grateful to all authors for their contributions and hope that this volume provides important insights into Macrocognition research, and a useful resource for research and application in this discipline. We are confident that Macrocognition has staying power, if only because of its complementarity to the traditional micro-cognitive paradigm.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct, and intellectual contribution to the work, and approved it for publication.

## REFERENCES

Crandall, B., Klein, G., and Hoffman, R. R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis.* Cambridge, MA: MIT Press.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., and Hollnagel, E. (2003). Macrocognition. *IEEE: Intell. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Schraagen, J. M., Militello, L. G., Ormerod, T., and Lipshitz, R. (eds.). (2008). *Naturalistic Decision Making and Macrocognition.* Aldershot: Ashgate.

# Designing System Reforms: Using a Systems Approach to Translate Incident Analyses into Prevention Strategies

*Natassia Goode\*, Gemma J. M. Read, Michelle R. H. van Mulken, Amanda Clacy and Paul M. Salmon*

*Faculty of Arts, Business and Law, Centre for Human Factors and Sociotechnical Systems, University of the Sunshine Coast, Maroochydore, QLD, Australia*

Advocates of systems thinking approaches argue that accident prevention strategies should focus on reforming the system rather than on fixing the "broken components." However, little guidance exists on how organizations can translate incident data into prevention strategies that address the systemic causes of accidents. This article describes and evaluates a series of systems thinking prevention strategies that were designed in response to the analysis of multiple incidents. The study was undertaken in the led outdoor activity (LOA) sector in Australia, which delivers supervised or instructed outdoor activities such as canyoning, sea kayaking, rock climbing and camping. The design process involved workshops with practitioners, and focussed on incident data analyzed using Rasmussen's AcciMap technique. A series of reflection points based on the systemic causes of accidents was used to guide the design process, and the AcciMap technique was used to represent the prevention strategies and the relationships between them, leading to the creation of PreventiMaps. An evaluation of the PreventiMaps revealed that all of them incorporated the core principles of the systems thinking approach and many proposed prevention strategies for improving vertical integration across the LOA system. However, the majority failed to address the migration of work practices and the erosion of risk controls. Overall, the findings suggest that the design process was partially successful in helping practitioners to translate incident data into prevention strategies that addressed the systemic causes of accidents; refinement of the design process is required to focus practitioners more on designing monitoring and feedback mechanisms to support decisions at the higher levels of the system.

Keywords: systems thinking, prevention strategies, learning, accidents, accident prevention

## INTRODUCTION

Incident reporting and investigation systems are now widely considered to be an essential component of safety management systems, and a pre-requisite for learning from incidents (Nielsen et al., 2006; Lindberg et al., 2010; Jacobsson's et al., 2011; Jacobsson et al., 2012). Most organizations have their own reporting and investigation systems; this is a requirement in the international standard for

occupational health and safety management (Nielsen et al., 2006). In safety critical domains, such as process control, aviation and healthcare, a number of sector-wide systems have existed since the early 1980s and 2000s (e.g., the Major Accidents Reporting System, 2012; Aviation Safety Reporting System, 2015; and the U.K.'s National Health Service Patient Safety reporting system, Department of Health, 2006). These sector-wide systems are intended to support cross-organizational learning from incidents, as well as reforms to regulation and legislation (Vincent, 2004; Jacobsson et al., 2010; Lindberg et al., 2010). Concerns have been raised, however, that there is little evidence that incident data is actually used to identify prevention strategies or support learning from incidents (Nielsen et al., 2006; Pless, 2008; Jacobsson et al., 2010; Lindberg et al., 2010). One of the reasons underpinning this is the absence of formal processes for translating incident data into appropriate accident prevention strategies[1]. This article describes and evaluates a new process for translating incident data analyses into prevention strategies, based on a systems thinking approach.

## Previous Research on Translating Incident Data into Prevention Strategies

For organizations, "learning from an incident" involves converting an incident experience into activities that will prevent future incidents (Jacobsson et al., 2012). Several models in the literature describe this process as a series of steps, where no one step can fail without affecting the end result (e.g., Lindberg et al., 2010; Jacobsson et al., 2012; Drupsteen et al., 2013a). Jacobsson, Ek and Akselsson (2011,?) "learning cycle" model describes the following steps: reporting; analysis; decision-making; implementation; and follow-up. "Reporting" includes the initial reporting and collecting additional data through investigation if required. "Analysis" describes the method for analyzing the data, and designing strategies that prevent similar incidents. "Decision-making" describes the process for selecting prevention strategies for implementation. "Implementation" describes the processes for converting the decisions into action. Finally, "Follow-up" includes both monitoring the implementation, and evaluating the impact of the action.

The majority of research examining aspects of the learning cycle has focused on the methods used to investigate incidents and analyze the data (for a review see Katsakiori et al., 2009). In addition, there is a significant body of research examining the factors influencing initial reporting, and the selection, implementation and maintenance of prevention strategies (e.g., Pidgeon and O'Leary, 2000; Lundberg et al., 2010, 2012; Ramanujam and Goodman, 2011; Le Coze, 2013; Vastveit et al., 2015). However, little research has focused on the process of designing prevention strategies, or describing the prevention strategies that result.

This lack of research into the design of prevention strategies implies that there is a belief that the analysis of incident data will *automatically* lead to new knowledge, new structures, new

---

[1]The term "prevention strategies" is used interchangeably in the literature with other terms such as "prevention strategies," "prevention activities," "recommendations," "remedial actions," "corrective actions," "countermeasures" and "interventions."

rules, and new practices that will result in higher reliability and improved safety once implemented (Lundberg et al., 2010; Carroll and Fahlbruch, 2011; Drupsteen et al., 2013b). However, examinations of investigation manuals show that little guidance is provided on how to design prevention strategies based on the outputs from an investigation (Lundberg et al., 2009; Rollenhagen et al., 2010; Drupsteen et al., 2013b). It is therefore unclear how safety practitioners design prevention strategies from the causes that are found, or prioritize addressing certain causes over others.

Another issue is that investigation manuals often give little consideration to understanding how the implementation of specific prevention strategies might impact on the system as a whole (Johnson, 2003; Lundberg et al., 2009; Rollenhagen et al., 2010). The approach to developing prevention strategies in many organizations is to address each cause identified in isolation (Johnson, 2003; Lundberg et al., 2009; Drupsteen and Hasle, 2014). This is problematic as changes to any system component will necessarily impact on others, and potentially lead to unintended, negative consequences (Lundberg et al., 2009; Kirwan, 2011). One reason for this may be that many investigations are still underpinned by linear chain-of-event accident causation models. These models focus safety practitioners on the negative events within an accident sequences and the "broken" components of the system. The underlying accident model therefore works against understanding the system as a whole (Lundberg et al., 2009; Rollenhagen et al., 2010; Dekker, 2011; Leveson, 2011).

A number of authors have argued that using a systems-based accident causation model to collect and analyze incident data might better support addressing problems holistically, rather than just treating individual parts of the system (Dekker, 2011; Leveson, 2011; Hollnagel, 2012). Systemic models are underpinned by three core principles of accident causation. First, safety in work systems is impacted by decisions and actions made at all levels of the system, not just by human operators working within the immediate context of the hazardous processes. Second, accidents are caused by multiple factors that go beyond the immediate context of the incident. Third, accidents and safety are described as emergent properties of systems, arising from interactions between the components within that system (Hollnagel, 2004; Leveson, 2011). Accidents and safety are considered to be "emergent properties" as the outcome of interactions between the components cannot be predicted from examining the functioning or reliability of each components in isolation (Dekker et al., 2011; Leveson, 2011). Based on these principles, it has been argued that prevention strategies should focus on addressing the factors at the higher levels of the system that create hazardous conditions and unsafe acts, rather than directly on failures relating to technology or human operators (e.g., Rasmussen, 1997; Dekker, 2011). In addition, it is the authors' opinion that these principles imply that organizations need to identify networks of prevention strategies, rather than standalone ones, in order to address failures arising from interactions between the components in the system.

A number of systems-based analysis methods have been developed that represent the contributing factors involved in accidents as complex, non-linear causal networks (e.g., STAMP,

Leveson, 2011; AcciMap, Rasmussen, 1997). Many studies have demonstrated that they provide a deeper understanding of how interactions within systems contribute to hazardous conditions and unsafe behavior in a range of safety-critical domains including space exploration (Johnson and Muniz de Almeida, 2008), aviation (Branford, 2011), rail (Underwood and Waterson, 2014), public health (Cassano-Piche et al., 2009), disaster management (Salmon et al., 2014a), road freight transport (Salmon et al., 2013; Newnam and Goode, 2015), and led outdoor activities (Salmon et al., 2014b, 2016a). Although these studies have focused on describing how accidents are caused, rather than how they can be prevented, there is no obvious reason why the same methods could not be applied to both analyze accidents and identify prevention strategies (Salmon et al., 2016b). Potentially, these methods could be extended to provide a structured process for translating incident data analyses into prevention strategies. If this approach is successful, the resulting prevention strategies should address the systemic causes of accidents.

This article investigates this proposition further by presenting the findings from a study using a systems approach to accident analysis and the prevention strategy design process. The study involved conducting participatory workshops with practitioners to identify prevention strategies from incident data collected through a national reporting system from the led outdoor activity (LOA) sector in Australia. The collection and analysis of the incident data, and the workshop prevention strategy design process, were all based on Rasmussen's (1997) risk management framework and associated AcciMap technique. The following sections provide a brief overview of both, along with details of their application to the LOA sector and the current study.

## Rasmussen's Risk Management Framework and AcciMap

Rasmussen's (1997) risk management framework is underpinned by the idea that work systems can be described as a hierarchy of multiple levels (e.g., government, regulators/associations, company, management, staff, work), as shown in **Figure 1**. The actions and decisions of those operating within and across these levels interact, and contribute to the control of hazardous processes. Safety is maintained through a process referred to as "vertical integration," where decisions made at higher levels of the system (i.e., by government, regulators, and the company) are reflected in practices occurring at lower levels of the system, while information at lower levels (i.e., work, staff) informs decisions and actions at the higher levels of the hierarchy. A lack of vertical integration can result in a loss of control and accidents (Svedung and Rasmussen, 2002; Cassano-Piche et al., 2009). The framework also describes how work practices constantly adapt and change in response to various external pressures and conditions. This process, referred to as "migration," causes accidents when changes in work practices erode existing control measures (Rasmussen, 1997).

The accompanying AcciMap technique provides a methodological framework for analyzing accidents from this perspective. The method enables analysts to graphically represent the contributing factors across all levels of the system in
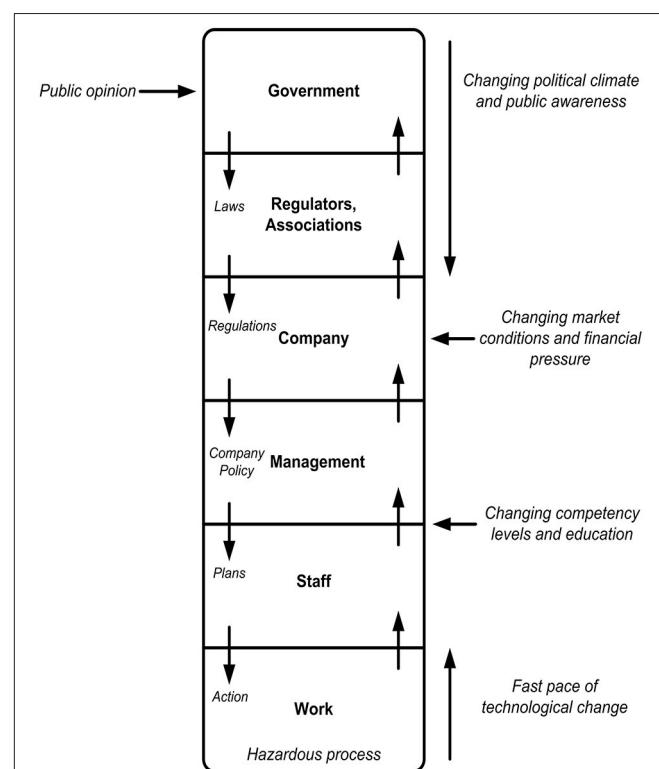


**FIGURE 1 | Rasmussen's risk management framework (adapted from Rasmussen, 1997).**

question, along with the relationships between them (Rasmussen, 1997; Svedung and Rasmussen, 2002).

Rasmussen's framework also makes a series of predictions, shown in **Table 1**, regarding accidents and safety in complex sociotechnical systems. These predictions reflect the three core principles of accident causation underpinning the systems approach, and also describe the role that vertical integration and the migration of work practices play in accident causation. These predictions have been used to evaluate the applicability of Rasmussen's framework and the AcciMap technique in new domains (e.g., Cassano-Piche et al., 2009; Jenkins et al., 2010; Salmon et al., 2014a), and to evaluate whether accident investigation processes adequately support the application of systems analysis methods (Newnam and Goode, 2015).

In the current study, the AcciMap technique was used initially to graphically represent the contributing factors, and the relationships between them, which were identified from incidents reported in the LOA sector in Australia. It was also subsequently used to represent networks of prevention strategies proposed to address these contributing factors and prevent future occurrences of similar incidents. Rasmussen's predictions were used to underpin the prevention strategy design process, and to evaluate whether the resulting prevention strategies address the systemic causes of accidents. These applications are described in detail in the following sections.

**TABLE 1 | Rasmussen's predictions regarding performance and safety in complex sociotechnical systems.**

| | |
|---|---|
| 1. | Safety is an emergent property—it is impacted by the decisions of all actors within the system |
| 2. | Accidents are caused by multiple contributing factors, not a single catastrophic decision or action |
| 3. | Accidents can result from a lack of vertical integration across levels, not just deficiencies at any one level alone |
| 4. | Lack of vertical integration is caused by a lack of feedback across levels. Actors cannot see how their decisions interact with those made by actors at other levels so threats to safety are not obvious before an accident |
| 5. | Work practices are not static, they migrate over time under the influence of a cost gradient driven by financial pressures in an aggressive competitive environment and under the influence of an effort gradient driven by the psychological pressure to follow the path of least resistance |
| 6. | Migration of work practices can occur at multiple levels, not just in one level alone |
| 7. | Migration of work practices causes the system's defenses to degrade and erode gradually over time |

## Application to Incident Data Collection and Analysis in the LOA Sector

The research described in this article was undertaken in the LOA sector in Australia. This sector includes all organizations that facilitate supervised or instructed "led" outdoor activities, such as outdoor education and recreation providers, school camps, adventure tourism operators and outdoor therapy programs (Goode et al., 2014a). These organizations deliver potentially high-risk activities (e.g., canyoning, sea kayaking, rock climbing, camping) in dynamic environments.

In the past 10 years, a number of high profile fatalities have occurred in Australia and internationally, which highlighted the need for better methods for understanding and preventing incidents in this domain (Salmon et al., 2010, 2012). For example, six students and their teacher died while on a gorge walking activity in New Zealand in 2008. The coroner and an independent investigation highlighted multiple contributing factors relating to the instructor, her manager, the activity center, the local weather service and government legislation and regulation (Brookes et al., 2009; Davenport, 2010). Previous literature on incident causation in this domain had focused on the immediate context of the incident (e.g., activity leader knowledge of environmental hazards and experience, supervision, weather), with little acknowledgment of the factors at the higher levels of the system (e.g., Curtis, 1995; Brookes, 2003, 2004).

There is now significant evidence that accident analysis methods underpinned by a systems approach are required to understand the incidents that occur during led outdoor activities. Analyses of fatal incidents (Salmon et al., 2010, 2012), near misses, and more common everyday injuries and illnesses (Salmon et al., 2014b, 2016a) have identified multiple contributing factors. In this domain, illnesses are viewed as important as even relatively minor illnesses or allergies may pose a serious risk in remote or wilderness environments (Goode et al., 2015).

To support the collection of incident data in the Australian LOA sector from a systems perspective, the authors have used Rasmussen's (1997) risk management framework to underpin the development of a national incident reporting system (Goode et al., 2015; Salmon et al., 2016a). The Understanding and Preventing Led Outdoor Accidents Data System (UPLOADS) allows organizations to record detailed information on incidents, including the event itself (e.g., the activity, the participants and supervisory staff involved), relevant events leading up the

incident, and describe the system of contributing factors that staff and management perceive to be involved. This data is then sent to the research team for analysis, and reports are produced annually.

To standardize the analysis of the incident data by the research team, the authors have developed a domain-specific contributing factor classification scheme, based on Rasmussen's framework and AcciMap technique. The classification scheme, shown in **Figure 2**, describes the actors and contributing factors involved in incidents across the LOA system. The classification scheme was developed and refined in a series of previous studies (Goode et al., 2014b; Salmon et al., 2014b; Taylor et al., 2015a,b).

Injury, illness and near miss incident data reported and analyzed via UPLOADS over a 12 month period (1st June 2014—31st May 2015) were used as the primary source of information for the prevention strategy development workshop. The prevention strategy design process focused on three AcciMaps representing the contributing factors identified from the injury, illness and near miss data. Due to space restrictions, only the prevention strategies relating to the injury data are presented in this paper.

## Application to the Prevention Strategy Design Process in This Study

Rasmussen's framework and the AcciMap technique were also used to underpin the prevention strategy design process. During the design process, the AcciMaps representing the incident data were used to identify specific goals for incident prevention. For each specific goal, a network of prevention strategies, and the potential relationships between them, were identified. Each prevention strategy identified a specific action and the actors that would be responsible for implementation. Relationships between the prevention strategies were used to describe how the successful implementation of one prevention strategy depended on another, or how the prevention strategies supported better vertical integration. The prevention strategies and the relationships between them were mapped onto the framework shown in **Figure 2** using the AcciMap technique (the resulting diagrams are referred to as PreventiMaps in this paper).

To guide the prevention strategy design process, Rasmussen's predictions were used to derive a series of reflection points (see **Table 2**). These reflection points were used by workshop facilitators to prompt practitioners to think about the incident data and prevention strategies from a systems perspective. In addition, a key question for this article was whether this design
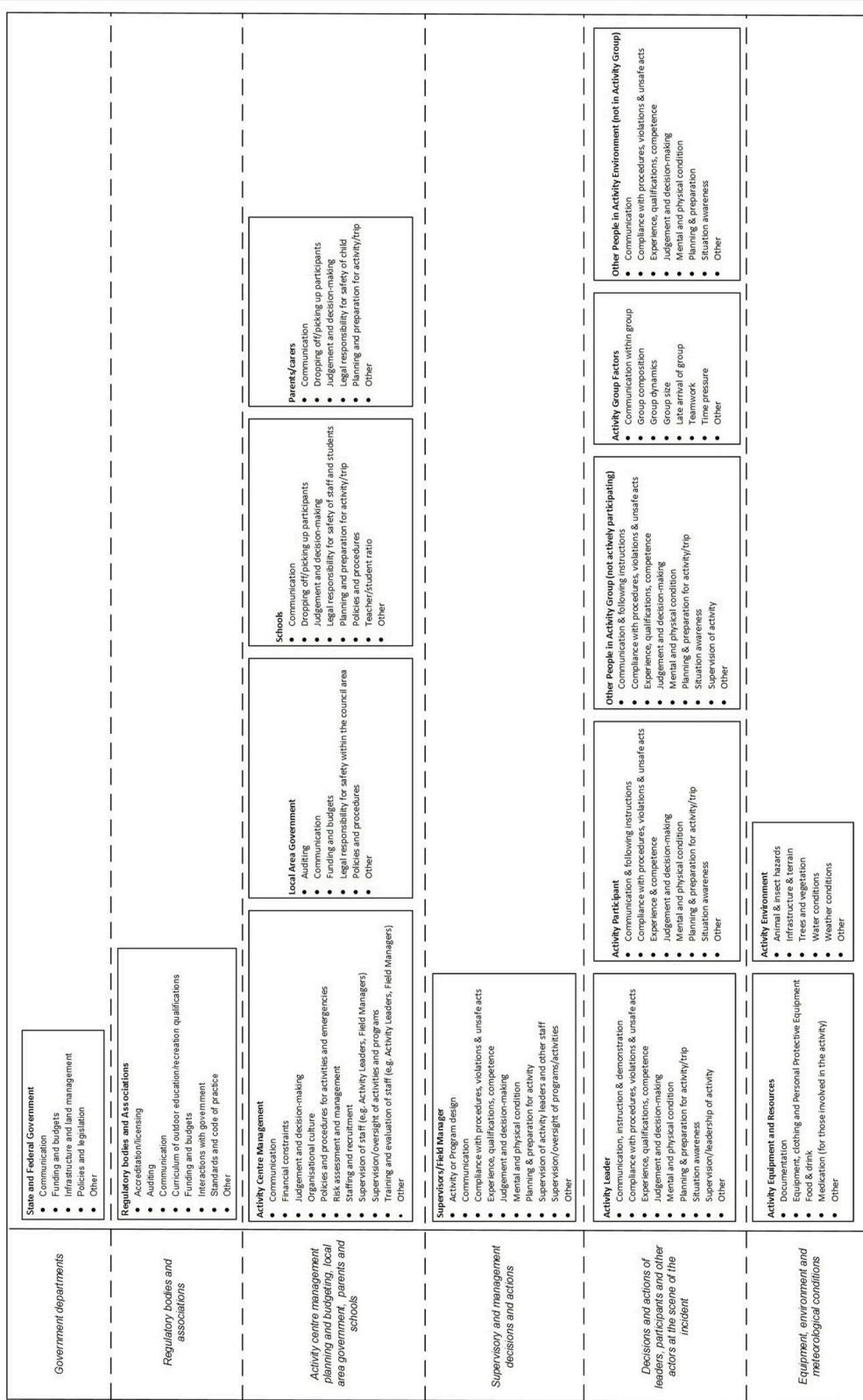
**FIGURE 2 | LOA contributing factor classification scheme based on Rasmussen's framework and AcciMap technique.**

process resulted in prevention strategies that addressed the systemic causes of accidents. Therefore, Rasmussen's predictions were also used to develop criteria for evaluating the networks of prevention strategies developed during the workshops (see **Table 2**).

In summary, the aims of this article are to: (1) describe the prevention strategies that were developed using a systems thinking approach; and (2) evaluate the extent to which they address the systemic causes of accidents as defined by Rasmussen's risk management framework.

## METHODS

### Design

Two workshops with practitioners from the LOA sector in Australia were conducted to design prevention strategies based on incident data. Ethics approval was obtained from the University of the Sunshine Coast Human Research Ethics Committee.

### Participants

Participants were invited to workshops based on their experience and role within the sector, or role in regulating safety within the sector. The aim was to ensure that the workshops included representatives from across the LOA system, including actors from the following: secondary schools; outdoor education providers; outdoor training organizations; outdoor sector Peak bodies; work health and safety (WHS) regulator; and relevant government departments.

In total, 30 people attended the workshops (Workshop 1 = 20, Workshop 2 = 10). The majority of participants were male (25 males, 5 females) and had a mean age of 47 years ($SD$ = 9.53), with a mean of 21 years' experience in the outdoor sector ($SD$ = 9.52, missing = 3). The number of workshop participants representing each actor within the sector is represented in **Figure 3** (note that some participants held more than one role).

### Workshop Planning Activities

Materials from a systems thinking-based design toolkit (Read et al., 2015), originally developed for use with the Cognitive Work Analysis (CWA) framework (Vicente, 1999), were adapted for use with the AcciMap analyses. The toolkit provides a structured approach for translating the outcomes of systems analysis methods into design concepts. The toolkit provided guidance on who should participate in the workshops and the type of group discussion activities required during the design process. Applying the toolkit resulted in a workshop plan and a set of reflection points to guide the design process based on Rasmussen's predictions (see **Table 2**).

### Materials

#### Incident Data and Analysis

The incident data was collected over a 12-month period (1st June 2014—31st May 2015) by 31 LOA organizations across Australia. The organizations used UPLOADS to collect information about the injuries, illnesses and near misses that occurred during LOA programs during this period. Injuries and illnesses were defined

as any issue that required care. This included any injury or illness requiring localized care with short term effects through to fatalities. A near miss was defined as "as a serious error or mishap that has the potential to cause an adverse event but fails to do so because of chance or because it is intercepted. For example, during a rock climbing activity an instructor notices that a participant's carabiener was not locked. If the student had fallen, this may have led to a serious injury." The organizations submitted deidentified data to the research team on a quarterly basis (van Mulken et al., 2016).

In total, 1020 incidents were reported, and 523 reports described the contributing factors and relationships involved in the incidents. These reports were analyzed by two members of the research team. This involved extracting a list of contributing factors and relationships between them from each report, discussing any discrepancies and reaching a consensus. The contributing factors and relationships were then classified using the scheme described in **Figure 2**. Summary AcciMaps were produced for each of the injury, illness and near miss data. This involved aggregating the contributing factor codes and the relationships between them across all the incidents within each type. The number of times the code and relationship appearing within the data were indicated on each AcciMap. Only the prevention strategies relating to the injury data are presented in this paper; **Figure 4** presents the AcciMap analysis for this data.

A report was then produced with sections on the injury, illness and near miss data. Each section of the report included descriptive statistics (e.g., led outdoor activities associated with incidents, severity ratings, demographics of people involved), AcciMaps, tables describing the specific contributing factors and relationships underpinning the information presented in the AcciMaps, and text descriptions of the findings.

For the workshop, summaries of the results were produced for the injury, illness and near miss data. In addition, large print-outs of the AcciMaps were given to each group, as well as blank AcciMap templates (i.e., diagrams with the six AcciMap levels labeled). These were used to document the networks of prevention strategies generated during the workshop.

### Procedure

Two workshops were held; one in Brisbane and one in Melbourne, Australia. Prior to the workshops, participants were emailed the aims of the workshop and the incident data report. The report was provided to give participants time to read through the analysis in detail.

On arrival at the workshop, participants were introduced to the objectives of the session and provided written consent to take part in the study. Participants were then presented with information about Rasmussen's risk management framework and the AcciMap method, and introduced to Rasmussen's predictions regarding accident causation. They were then given a presentation on the key findings from the analysis of the injury, near miss and illness data, including an overview of the AcciMaps. They were given instructions on how to interpret the AcciMaps and data tables within the report and were given an example of why component-orientated prevention strategies might be unsuccessful. They were also provided

**TABLE 2 | Reflection points developed for the prevention strategy design process and the criteria used to evaluate the resulting PreventiMaps based on Rasmussen's predictions.**

| | Reflection points | Evaluation criteria |
|---|---|---|
| 1 | Can you see in the AcciMap how decisions and interactions between actors created situations where incidents occurred? | The prevention strategies require actions and decisions from multiple actors (at least three). |
| | This solution relates to one actor, can you think of related solutions that fit in other levels of the AcciMap structure? | The prevention strategies require changes at multiple levels of the system (at least three). |
| | How does the solution support interaction/coordination across actors at different levels? | |
| 2 | Is there an obvious set of contributing factors in the AcciMap that appears to be important? | Multiple interdependent prevention strategies are identified to address the specified goal (at least three). These include mechanisms to support the implementation of prevention strategies within and across levels. |
| | Could this solution be part of a wider set—what is needed at the level above to make it work? What is needed at the level below? | |
| 3 | Can we improve communication and coordination across the levels to improve this issue? | The prevention strategies support the flow of information from actors across and within system levels. |
| | Could information flowing upwards be improved? | |
| | Could information flowing downwards be improved? | |
| | Could information flow within actors at the same level be improved? | |
| | To make this idea work what would need to be communicated up to the higher levels? What would need to be communicated down to the lower levels? | |
| | To make this solution work, how does information need to flow between actors—upwards, downwards, and across levels of the system? | |
| 4 | Can we improve feedback across levels of the system so that an actor knows the outcomes of their decisions and actions? | The prevention strategies improve feedback processes to actors regarding the impact of their decisions and actions. |
| 5 | How might financial pressures impact on this solution, especially over time? Is it financially sustainable? Can we improve this? | The prevention strategies provide mechanisms for actors at the higher levels to identify or monitor changes to work practices at the frontline of operation. |
| | How might psychological pressures impact on this solution, especially over time? Will people see its ongoing relevance? Can we improve this? | |
| | How could we identify or monitor changes to work practices as a result of financial pressures or psychological pressures? | |
| 6 | How might financial pressures at a higher/lower level of the system impact on this solution? | The prevention strategies provide mechanisms for monitoring changes to work practices for actors at the higher levels of the system. |
| | How might psychological pressures at a higher/lower level of the system impact on this solution? | |
| 7 | How could we monitor whether defenses are degrading/eroding over time within organizations and/or across the sector? | The prevention strategies include mechanisms for monitoring whether the implementation of risk control measures are degrading over time. |

*Numbers relate to the predictions shown in* **Table 1**.

with a simple example of a network of prevention strategies relating to the prevention of blisters, mapped onto an AcciMap template.

Next, participants partook in small group discussions, with each group led by a facilitator. These discussions were audio-recorded using a dictaphone. The discussions occurred in three rounds, each lasting approximately 45 min each. In the first round, participants considered the injury data, in the second round the illness data and in the third round the near miss data. Participants remained in the same small group for each round. There was a total of 7 groups across both workshops.

At the start of each round of discussion, participants were first asked to review the AcciMaps and data tables, and discuss the contributing factors. Where participants offered additional contributing factors that they believed from

experience had a role in the events, these were documented by the facilitator. Participants were then encouraged to discuss potential prevention strategies and to consider how prevention strategies could be linked in a network or cluster of prevention strategies across the LOA system. Participants could choose whether to focus on developing prevention strategies to address specific issues identified in the data (e.g., burns resulting from cooking and campfires), or the total dataset. The reflection points were used either to prompt initial ideas or to refine ideas that were generated by participants. The facilitators documented the prevention strategies, and links between them, on the blank AcciMap templates. Each prevention strategy was described on the AcciMap in terms of the actors primarily responsible for implementation and the specific actions required (e.g., "National Parks: change camping permits to improve access to severe
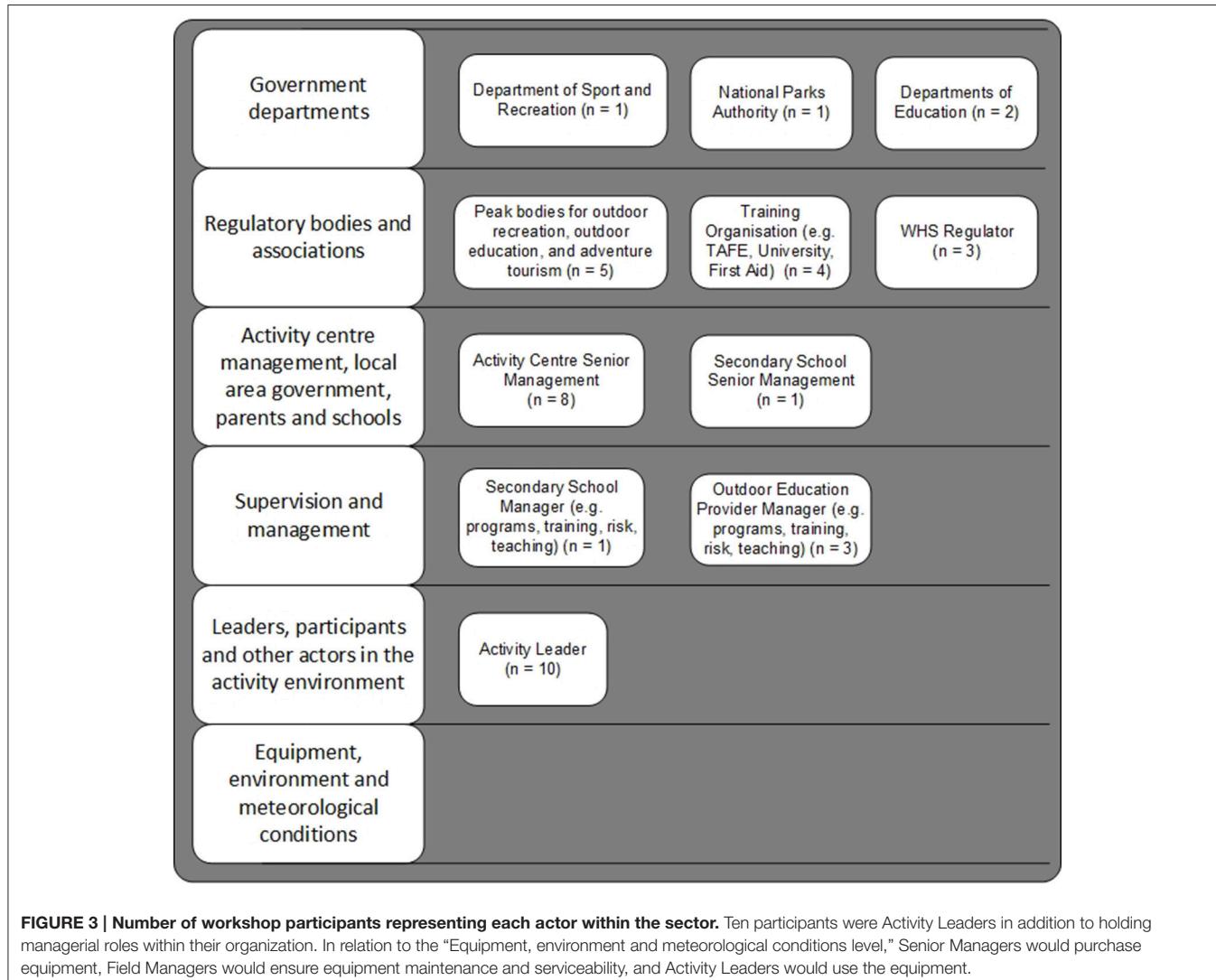
**FIGURE 3 | Number of workshop participants representing each actor within the sector.** Ten participants were Activity Leaders in addition to holding managerial roles within their organization. In relation to the "Equipment, environment and meteorological conditions level," Senior Managers would purchase equipment, Field Managers would ensure equipment maintenance and serviceability, and Activity Leaders would use the equipment.

weather camping sites when required"). At the conclusion of the discussions, the facilitators presented each PreventiMap to the group, and made any additions or changes based on feedback.

## Data Analysis

Due to space restrictions, only the prevention strategies relating to the injury data were analyzed for this paper.

The 7 PreventiMaps developed by the groups to address the key findings from the injury data were represented in Microsoft Visio. Each PreventiMap was reviewed and amended (to ensure clarity of description) by the facilitator who had originally documented it. Audio recordings were used when further information was needed to provide a more specific description of the prevention strategies. In addition, where appropriate, the facilitator created separate PreventiMaps to represent the specific goals their groups had discussed. This resulted in 10 PreventiMaps representing specific goals for incident prevention based on the injury data.

To identify similar prevention strategies across the groups, the PreventiMaps were coded using Nvivo 10. Each individual prevention strategy was coded into a theme based on: (1) the actors identified as responsible for implementation (e.g., Peak body); and (2) the specific actions required (e.g., lobby the government regarding the need to educate community on the benefits of LOA). A summary PreventiMap was then constructed by the researchers representing the prevention strategies that were identified by the workshop groups.

In addition, the 10 PreventiMaps representing specific goals for incident prevention were evaluated using the criteria presented in **Table 2**. The evaluation involved examining each PreventiMap, and giving a "Yes," "Partial," or "No" rating based on the criteria. "Yes" and "Partial" ratings had to be supported by examples, which were recorded in a table. The evaluation was initially conducted by the first author, and then validated by the second author. Any disagreements were resolved through discussion.
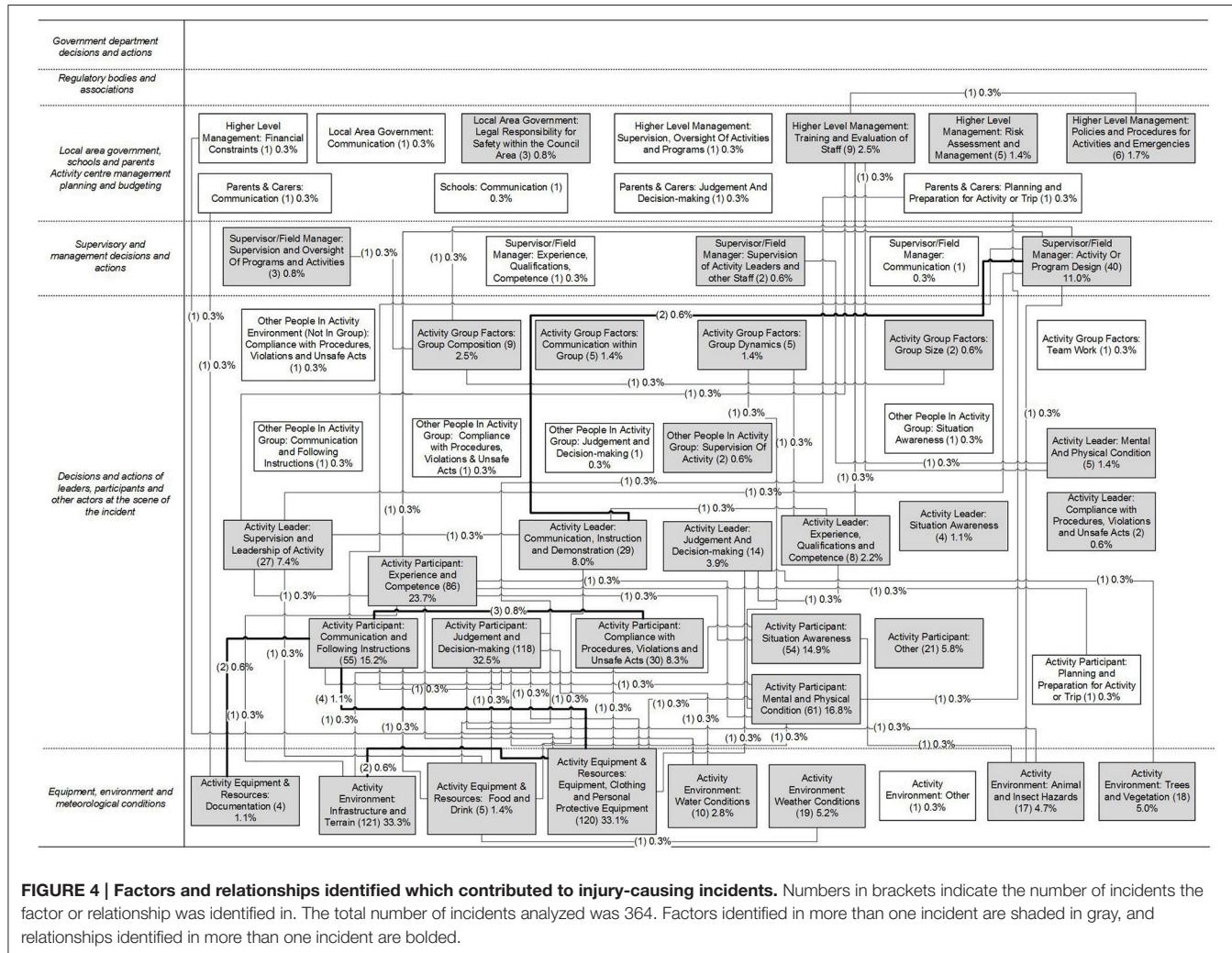
**FIGURE 4 | Factors and relationships identified which contributed to injury-causing incidents.** Numbers in brackets indicate the number of incidents the factor or relationship was identified in. The total number of incidents analyzed was 364. Factors identified in more than one incident are shaded in gray, and relationships identified in more than one incident are bolded.

## RESULTS

This section first presents an overall summary of all the prevention strategies identified by the workshop groups in relation to the injury data, as well as an example of a PreventiMap developed to address a specific goal. A summary of the findings from the evaluation is then presented. Note that throughout the results section "n" refers to the number of workshop groups (total $n = 7$).

## Description of Prevention Strategies

Based on the injury data, the workshop groups identified the following specific goals for incident prevention:

1. The prevention and management of Activity Leader fatigue (Group 1)
2. The prevention of burns during cooking activities (Group 2)
3. Improvement of participants' skills for outdoor activities (Group 3)
4. Improvement of reporting of pre-existing injuries (Group 3)
5. Ensuring that the difficulty of program matches participants' competence level (Group 4)

6. Improvement of communication around participant competence levels (Group 5)
7. Improvement of participants' physical literacy (Group 5)
8. Improvement of activity leaders' competencies around dynamic risk assessment (Group 6)
9. Professionalization of the career pathway for people in the LOA sector (Group 6)
10. Improvement of activity leaders' competencies for dealing with injuries (Group 7).

**Figure 5** shows a summary of all the prevention strategies identified to address these goals. Notably, prevention strategies were identified at all levels of the LOA system and in relation to all actors represented within the UPLOADS classification scheme. Some prevention strategies specifically addressed improving communication and collaboration between actors. The majority of prevention strategies focused on actions required at the second and third level of the framework. The actors most frequently identified as responsible for implementation were Peak bodies and Activity Center Management. The prevention strategy themes most frequently identified were "Peak bodies:

**FIGURE 5 | Summary of the prevention strategies identified by workshop participants in relation to the injury data, presented according to the actors responsible for implementing the prevention strategy and the key themes.** Numbers in brackets indicate the number of workshop groups that identified the theme. The total number of workshop groups was 7.

Changes to policies and standards" ($n = 6$), "Activity Center: Improve communication and coordination between Activity Centers, schools and parents" ($n = 5$), and "Activity Center: Provision of training for Activity Leaders ($n = 6$).

All prevention strategies that were coded as "Peak bodies: Changes to policies and standards" focused on changes to the Adventure Activity Standards (AAS), which are voluntary safety guidelines for organizations conducting LOA. For example, to improve the quality of supervision during programs, Group 1 suggested that the AAS should "...incorporate Activity Leaders hours of work spent driving, active supervision and inactive supervision during programs," while Group 2 suggested that the AAS should "...include supervision requirements and ratios around camp craft and camping." Both prevention strategies were in response to the finding that "Activity Leader: Supervision and Leadership of Activity" and "Activity Leader: Communication, Instruction and Demonstration" were involved in just under 10% of all injury-causing incidents as shown in **Figure 4**.

The majority of prevention strategies coded as "Activity Center: Improve communication and coordination between Activity Centers, schools, and parents" focused on improving communication regarding participant experience, abilities and pre-existing injuries. For example, Group 2 suggested that Activity Centers should "improve communication with parents about child's previous experience outdoors," while Group 6 suggested they should improve "...communication between

schools and Activity Centers around participants health and abilities." These prevention strategies were in response to the finding that many injury-causing incidents were caused by "Activity Participant: Experience and Competence" and "Activity Participant: Mental and Physical Condition," which were identified in 24% and 17% of injury-causing incidents, respectively, as shown in **Figure 4**.

The prevention strategies coded as "Activity Center: Provision of training for Activity Leaders" addressed a range of weaknesses discussed in relation to Activity Leader skill sets. For example, Group 1 suggested that Activity Centers should "...provide soft skill training for co-leaders and distributed leadership," while Group 4 suggested "...training for instructors to assist them to adapt program designs to suit the competence of the group." Again, these prevention strategies were in response to a range of contributing factors relating to Activity Leaders supervision, competence and decision-making, as well as the incidents involving issues with Activity or Program design (identified as a contributing factor in 7% of injury-causing incidents, shown in **Figure 4**).

## Example of a PreventiMap

**Figure 6** shows an example of the PreventiMaps developed by Group 4 to "ensure that the difficulty of the program matches participant skill levels." This was in response to two of the most frequently identified contributing factors in injury-causing incidents: "Activity Participant: Experience and
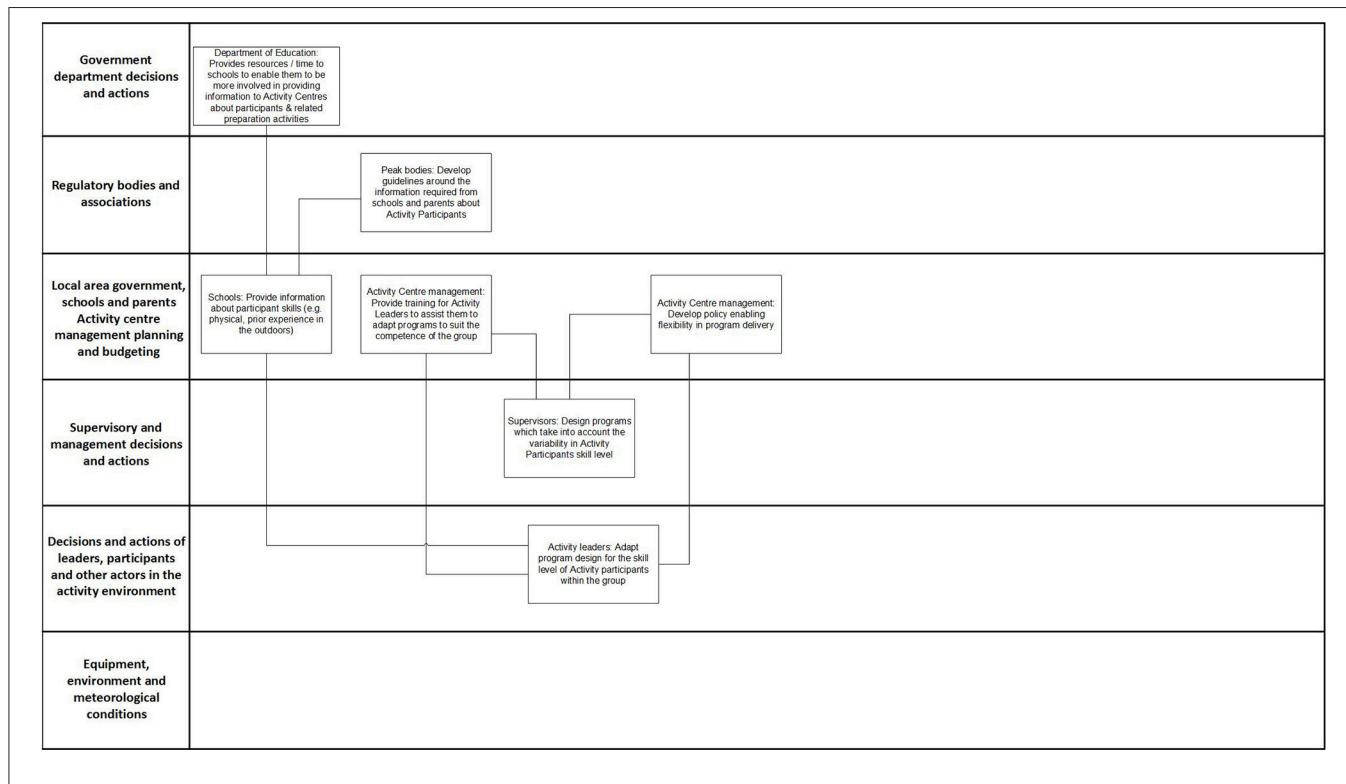
**FIGURE 6 | PreventiMap developed by Group 4 to ensure that the difficulty of program matches participants' competence level.**

Competence" and "Activity Participant: Communication and Following Instructions." These factors were identified in 24% and 15% of the injury incidents, respectively, and were highly interconnected to other factors on the AcciMap (see **Figure 4**). Workshop participants believed that many injuries occurred because program design did not adequately take into account Activity Participants' level of experience in the outdoors, and Activity Participants were ill prepared for the program (in terms of both physical literacy/fitness and equipment). Workshop participants discussed their perception that the skill level of participants had decreased over time, as children were less exposed to the outdoors and physical activity in their daily lives than previously.

The prevention strategies focus on improving communication between different actors within the system regarding participants' skills and implementing systems to increase the flexibility of program design. For example, workshop participants suggested that the Department of Education should provide more resources and time to enable schools to prepare participants for programs and gather information about their skills and abilities, which in turn, would enable schools to collect and provide information to Activity Centers on participants' competence. Activity Centers would then feed this information down into the development of programs. Workshop participants also suggested that Activity Leaders should be able to dynamically adapt programs to suit the skills of the group. They suggested that training on how to identify the skills of participants and adapt programs, as well as specific policies enabling flexibility in program delivery,

would be needed to support Activity Leaders performing this function.

## Evaluation of PreventiMaps

The evaluation focused on the 10 PreventiMaps representing specific goals for incident prevention (described in Section Description of Prevention Strategies). The following sections present the findings in relation to the criteria, with selected examples to support the ratings. The PreventiMaps are referred to by the numbers shown in **Table 3**, which also summarizes the ratings from the evaluation. **Table 4** summarizes the findings supporting the ratings for the first three evaluation criteria.

### Criterion 1: The Prevention Strategies Require Actions and Decisions from Multiple Actors (at Least Three)

All 10 PreventiMaps met this criterion. The PreventiMaps identified between 4 and 7 actors responsible for implementation. The actors most frequently identified as responsible were Peak bodies and Activity Center Management. While many of the contributing factors in the incident data related to Activity Participants, only one prevention strategy identified Activity Participants as playing a role in implementation.

### Criterion 2: The Prevention Strategies Require Changes at Multiple Levels of the System (at Least Three)

All 10 PreventiMaps met this criterion. The PreventiMaps required changes to 3–5 system levels. All PreventiMaps required

TABLE 3 | Summary of evaluation ratings for each PreventiMap representing specific goals for incident prevention.

| PreventiMap | Goal | Criteria | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | The prevention and management of Activity Leader fatigue | Yes | Yes | Yes | Partial | No | No | No | No |
| 2 | The prevention of burns during cooking activities | Yes | Yes | Yes | Partial | No | No | No | No |
| 3 | Improvement of participants' skills for outdoor activities | Yes | Yes | Yes | Yes | No | No | No | No |
| 4 | Improvement of reporting of pre-existing injuries | Yes | Yes | Yes | Yes | No | Yes | No | Yes |
| 5 | Ensuring that the difficulty of program matches participants competence level | Yes | Yes | Yes | Yes | No | No | No | No |
| 6 | Improvement of communication around participant competence levels | Yes | Yes | Yes | Yes | No | No | Partial | No |
| 7 | Improvement of participants' physical literacy | Yes | Yes | Yes | Partial | No | Partial | No | No |
| 8 | Improvement of activity leaders' competencies around dynamic risk assessment | Yes | Yes | Yes | Yes | No | Yes | No | No |
| 9 | Professionalization of the career pathway for people in the LOA sector | Yes | Yes | Yes | Yes | No | Partial | Partial | No |
| 10 | Improvement of activity leaders' competencies for dealing with injuries | Yes | Yes | Yes | Yes | No | Yes | Partial | Yes |

TABLE 4 | Summary of the findings supporting the ratings for the first three evaluation criteria.

| PreventiMap | Criterion 1 — Actors identified as responsible for implementation | | | | | | | | | | | | Criterion 2 — LOA system levels required to change | | | | | | Criterion 3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Government | Department of Education | Universities/TAFEs | Research groups | Peak bodies | Schools | Parents | Activity Center Management | Supervisors/Field Managers | Activity Leaders | Activity Participants | | 1 | 2 | 3 | 4 | 5 | 6 | No. of prevention strategies | No. of relationships |
| 1 | | X | X | | X | X | | X | X | | | | X | X | X | X | | | 13 | 15 |
| 2 | | X | X | | X | X | | X | X | X | | | X | X | X | X | X | | 13 | 8 |
| 3 | X | | | | X | X | X | X | X | X | X | | X | | X | X | X | | 9 | 7 |
| 4 | | X | | | X | | X | X | | | | | X | X | X | | | X | 6 | 6 |
| 5 | | X | | | X | X | | X | | X | | | X | X | X | | X | | 7 | 7 |
| 6 | X | X | | | X | X | X | X | | | | | X | X | X | | | | 16 | 20 |
| 7 | X | X | | | X | X | X | | | | | | X | X | X | | | | 6 | 9 |
| 8 | X | | X | | X | | | X | | | | | X | X | X | | | | 9 | 9 |
| 9 | | | X | X | X | | | X | | | | | X | X | X | | | | 6 | 4 |
| 10 | | | | | X | X | | X | X | X | | | X | X | X | X | | | 11 | 19 |

*LOA system levels: 1, Government department decisions and actions; 2, Regulatory bodies and associations; 3, Local area government, schools and parents, Activity Center management planning and budgeting; 4, Supervisory and management decisions and actions; 5, Decisions and actions of leaders, participants and other actors in the activity environment; 6, Equipment, environment and meteorological conditions.*

changes at the third level of the framework, and overall, they tended to focus on changes at the three highest levels of the system.

## Criterion 3: Multiple Interdependent Prevention Strategies Are Identified to Address the Specified Goal (at Least Three). These Include Mechanisms to Support the Implementation of Prevention Strategies within and across Levels

All 10 PreventiMaps met this criterion. The PreventiMaps described between 6 and 16 prevention strategies, and 4–20 relationships.

Most of the mechanisms identified to support implementation occurred across levels. For example, a number of prevention strategies at the higher levels were identified to support the prevention strategy: "Activity Leaders adapt program design for their group," including: flexibility is included in program design; Activity Centers provide training on how to adapt programs to suit competence levels; and Activity Centers develop a policy allowing Activity Leaders to change the delivery of programs (PreventiMap 5).

PreventiMap 3 included examples of across level support mechanisms. The prevention strategy "Activity Centers and Schools improve communication with parents around participant capabilities" was supported within the level by: schools improve briefing to parents around required levels of competence; and Activity Centers develop key descriptors of competence related to different types of activities.

## Criterion 4: The Prevention Strategies Support the Flow of Information from Actors Across and within System Levels

Seven of the PreventiMaps fully met this criterion, and three partially met this criterion.

The PreventiMaps that fully met this criterion included prevention strategies to improve the flow of information between actors both across and within levels. For example, PreventiMap 8 included prevention strategies to improve communication across and within levels regarding risk assessments. Specifically, workshop participants identified the prevention strategy "Peak bodies provide opportunities to talk with Activity Centers about risk assessment and share experiences" to improve across level communication, while "Activity providers to provide risk assessments to parents, and consent forms are signed based on this information" was identified to improve within level communication.

The PreventiMaps that partially met this criterion only included prevention strategies that increased the flow of information in a specific direction. For example, PreventiMap 7 only targeted the flow of information between Level 1 and Level 2 of the LOA system (e.g., "Peak bodies to lobby government to establish independent body on physical literacy").

## Criterion 5: The Prevention Strategies Improve Feedback Processes to Actors Regarding the Impact of Their Decisions and Actions

None of the PreventiMaps met this criterion. During the evaluation, it was noted that many of the PreventiMaps failed to identify mechanisms to monitor the impact of changes to regulations, policies and procedures. For example, although PreventiMap 1 describes a range of regulations, policies, and programs to prevent Activity Leader fatigue, no mechanism was identified for monitoring actual levels of fatigue.

## Criterion 6: The Prevention Strategies Provide Mechanisms for Actors at the Higher Levels to Identify or Monitor Changes to Work Practices at the Frontline of Operation

Three of the PreventiMaps fully met this criterion, and two met it partially.

The PreventiMaps that fully met this criterion included prevention strategies to monitor changes to Activity Leader work practices. For example, PreventiMap 10 included a prevention strategy specifying that Activity Leaders should receive training on "understanding and identifying complexities of mental and physical health issues." To monitor the impact of this program, it was proposed that Activity Centers should conduct "regular appraisals by peers and management to assess performance strengths and weaknesses to guide additional training."

The PreventiMaps that partially met this criterion only implied avenues for monitoring changes to work practices at the frontline of operation. For example, PreventiMap 7 proposed that the government should provide more funding for school outdoor education programs, and change the school curriculum to include outdoor education. Potentially, government departments would monitor the take up of this funding and the implementation of changes to school curriculum; however, this was not explicitly specified by participants.

## Criterion 7: The Prevention Strategies Provide Mechanisms for Monitoring Changes to Work Practices for Actors at the Higher Levels of the System

None of the PreventiMaps fully met this criterion, and three met it partially.

The PreventiMaps that partially met this criterion only implied avenues for monitoring changes to work practices at the higher levels of the system. For example, PreventiMap 7 included a relationship between "Peak bodies to lobby government to establish independent body on physical literacy" and "Government to increase funding for outdoor education programs." Potentially, the Peak bodies would monitor changes to funding at the government level, although this is not explicit. Similarly, PreventiMap 10 specified that "Activity Centres should set guidelines around the required number of permanent staff" to address the issues identified with causal staff lacking relevant knowledge and training. This prevention strategy would potentially prevent Activity Centers from hiring more causal staff in response to financial pressures.

## Criterion 8: The Prevention Strategies Include Mechanisms for Monitoring Whether the Implementation of Risk Control Measures Are Degrading Over Time

Two of the PreventiMaps met this criterion. For example, the goal of PreventiMap 4 was to improve the reporting of pre-existing injuries. It was proposed that "data on incidents rates are made available on the websites of Peak bodies." This provides a way of monitoring whether the risk control measures associated with reporting pre-existing are eroding over time at an industry level. Similarly, the goal of PreventiMap 10 was to improve activity leaders' competencies for dealing with injuries. It was proposed that activity leaders should receive "Regular appraisals by peers and management to assess performance." This provides a way for organizations to monitor whether the risk control measures associated with dealing with pre-existing injuries are eroding over time.

In relation to this criterion, it was noted during the evaluation that some of the prevention strategies proposed might have the unintended consequence of eroding risk control measures over time. For example, PreventiMap 5 focused on increasing the flexibility of the delivery of programs, with the expectation that Activity Leaders would alter programs to match Activity Participants level of competence. However, Activity Leaders might become more focused on altering programs than ensuring that existing risk controls are maintained. In addition, altering programs might unintentionally result in new hazards. No prevention strategies were proposed to address these potential consequences.

# DISCUSSION

The aims of this article were to describe the prevention strategies that were developed by applying a systems thinking approach during the design process, and to evaluate the extent to which they addressed the systemic causes of accidents as defined by Rasmussen's risk management framework. Using a systems thinking-based design process, workshop groups identified a range of specific goals for incident prevention from the injury data. To address these goals, PreventiMaps were developed representing prevention strategies requiring actions from all actors, across all levels of the system. All of the PreventiMaps required actions at the higher levels of the system, and only a few focused on the immediate context of LOA delivery. Prevention strategies involving actions at the frontline of system operation (e.g., Activity Leaders should adapt programs to suit participants capabilities) were supported by changes to policies, training and regulation. The subsequent evaluation of the PreventiMaps revealed that all of them addressed the three core principles of the systems approach (Criteria 1, 2, and 3), and the majority proposed prevention strategies for improving vertical integration (Criterion 4). However, overall the PreventiMaps tended to focus on top-down controls, rather than bottom-up feedback and monitoring of work practices. Therefore, the majority of the PreventiMaps failed to address Rasmussen's predictions regarding the migration of work practices over time and the erosion of risk control measures (Criterion 5–8). Overall, the evaluation shows that the design process was partially successful in helping practitioners to translate incident data into prevention strategies that address the systemic causes of accidents, and highlights areas for improvement in the design process.

The findings from this study will be used to improve the design process in a number of ways. First, the findings indicate that the reflection points need to be refined to focus practitioners more on identifying ways to monitor behavior and decision-making at the frontline of system operation, and designing feedback mechanisms to support decisions at the higher levels of the system. To support this aspect of the design process, it might be helpful to identify specific incidents from the LOA data where monitoring and feedback processes have failed, along with examples of successful monitoring and feedback processes used in other safety-critical domains. Second, the design process resulted in many, often overlapping, specific goals for incident prevention and prevention strategies. A further phase in the design process is required to refine and select specific goals and prevention strategies. This will require the development of further evaluation criteria to assist this decision-making process.

The approach used to design prevention strategies in this paper resulted in different outputs to the component-orientated approaches described in the literature (Johnson, 2003; Hollnagel, 2008; Lundberg et al., 2009, 2010; Rollenhagen, 2011). For example, based on an analysis of investigation manuals, Johnson (2003) describes four general approaches that are used by organizations to generate possible prevention strategies: the perfectibility approach; the heuristic approach; navigational techniques; and accident prevention models such as Haddon's (1980; see pp. 565–590 of Johnson, 2003, for a description

and extensive discussion of the strengths and weaknesses of each approach). The key difference between the approaches is the type of "fixes" that are deemed appropriate or effective. However, all of the approaches focus on developing a list of prevention strategies; each item on the list is intended to address a specific component of the problem identified in the incident analysis. No consideration is given to the relationships between prevention strategies or the interactions between them once implemented. The approach used in this study allowed participants to understand the interdependencies between the solutions they were proposing, and to identify the mechanisms needed to support implementation across the system.

It should be acknowledged that the group problem solving approach to designing prevention strategies is not novel, although it does not appear to be consistently used across industries (Lundberg et al., 2010; Rollenhagen et al., 2010; Rollenhagen, 2011). In the context of Swedish nuclear power plants, Rollenhagen (2011) notes that problem solving groups that include representatives from the whole system of interest are perceived as more successful in identifying more effective prevention strategies. He attributes the perception of success to increasing actors understanding of the system functions they do not directly influence, and of the consequences of their decisions for other functions. The key difference between earlier studies and the current study is the boundary on the "system of interest." In this study, the system included actors outside the context of the organization (e.g., Peak bodies, WHS regulators, and government departments). These actors had a detailed understanding of the guidelines, regulations, government policies, programs, and funding influencing LOA provision. This knowledge may have helped representatives from LOA providers think outside the "silo" of their organization, and consider the sector as a whole.

An unaddressed question from this study is the practicality of the prevention strategies proposed and whether the prevention strategies are likely to be implemented by the sector. For example, resource constraints are typically a significant factor that moderate the success of the prevention strategies proposed in response to incidents (Lundberg et al., 2010). The next phase of the research program involves inviting the whole sector to evaluate of the feasibility of the prevention strategies. However, it is acknowledged that even if the prevention strategies are favorably assessed by the sector, there are many factors that will influence their implementation (e.g., Pidgeon and O'Leary, 2000; Lundberg et al., 2010, 2012; Ramanujam and Goodman, 2011; Le Coze, 2013; Vastveit et al., 2015). A direction for future research is to chart the barriers to implementing system reforms that exists within the LOA sector.

The limitations of the study should be acknowledged, which also present some directions for future research. One significant limitation of the present study is a lack of comparison groups. For example, until the implementation of UPLOADS, the sector did not have good quality incident data to focus their preventative efforts (Goode et al., 2014a, 2015; Salmon et al., 2016a). Therefore, the same prevention strategies may have been identified based on the incident data analysis, without the design process. However, it seems unlikely that the networks of prevention strategies would have been generated without

the design process or application of the AcciMap technique. To address this issue, the authors plan to conduct controlled trials to compare the design process against unstructured group brainstorming sessions. In addition to evaluating the extent to which the proposed prevention strategies address the systemic causes of accidents, a scale developed by Jacobsson's et al. (2011) will be used to evaluate their potential effectiveness. This scale evaluates the effectiveness of prevention strategies on three dimensions: geographical application, degree of organizational learning, and time. More effective prevention strategies are those that apply across the organization, target the redesign of organizational systems, and involve plans for long term maintenance. The authors also plan to evaluate potential improvements in prevention strategies, and modifications to the design process, as the process is implemented in an organization over an extended period of time. Future studies are also required to determine the training requirements of implementing the design process in an organization to ensure that it produces valid outputs (Stanton and Stevenage, 1998; Stanton, 2016).

A second limitation of this study was that two important actors were missing from the workshops—activity participants and the parents of children involved in the activities—the LOA sector's "consumers." These actors may have a different view on the factors that would encourage them to play a more active role in managing risk. Accordingly, it is recommended that they are represented at future workshops.

In conclusion, the approach applied in this study allowed practitioners to create networks of prevention strategies designed to address the conditions at the higher levels of the LOA system. This approach to prevention strategy design is not only novel for the LOA sector, but across the safety critical domains. To the authors' knowledge, this study is the first reported to apply Rasmussen's Risk Management Framework and AcciMap technique to incident data collection, incident analysis and prevention strategy design, all as part of an integrated process. Most importantly, the prevention strategies were designed by the actors within the system of interest, rather than by researchers studying the system. We encourage further applications of the approach, and future research should consider how these methods might apply to the next steps in the learning cycle

(Jacobsson et al., 2012): decision-making, implementation and follow-up.

## AUTHOR CONTRIBUTIONS

NG Conceived and designed of the study, organized the data collection, facilitated the workshops, conducted the evaluation, analyzed the data and wrote the manuscript. GR Contributed to the conception and design of the study, facilitated the workshops, conducted the evaluation, contributed to writing and revising the manuscript. MV Contributing to organizing the data collection, facilitated the workshops, transcribed the data, contributed to analyzing the data, and contributed to revising the manuscript. AC Facilitated the workshops, transcribed the data, contributed to analyzing the data, and contributed to revising the manuscript. PS Conceived and designed of the study, facilitated the workshops, contributed to the evaluation, and contributed to writing and revising the manuscript.

## ACKNOWLEDGMENTS

## REFERENCES

Branford, K. (2011). Seeing the big picture of mishaps: applying the AcciMap approach to analyze system accidents. *Aviat. Psychol. Appl. Hum. Factors* 1, 31–37. doi: 10.1027/2192-0923/a00005

Brookes, A. (2003). Outdoor education fatalities in Australia 1960-2002. Part 2. Contributing circumstances: supervision, first aid, and rescue. *Aus. J. Outdoor Educ.* 7, 34–42. Available online at: http://www.latrobe.edu.au/education/downloads/brookes_a_OE-Fatalities2.pdf (Accessed December 15, 2016).

Brookes, A. (2004). Outdoor education fatalities in Australia 1960-2002. Part 3. Environmental circumstances. *Aus. J. Outdoor Educ.* 8, 44–56. Available online at: http://www.latrobe.edu.au/education/downloads/brookes_a_OE-Fatalities3.pdf (Accessed December 15, 2016).

Brookes, A., Smith, M., and Corkill, B. (2009). *Report to the Trustees of the Sir Edmund Hillary Outdoor Pursuit Centre of New Zealand: Mangatepopo*

*Gorge Incident*. Available online at: http://www.hillaryoutdoors.co.nz/newsite/wp-content/uploads/2013/06/091015-IRT-OPC_-Report.pdf

Carroll, J. S., and Fahlbruch, B. (2011). The gift of failure: new approaches to analyzing and learning from events and near-misses. Honoring the contributions of Bernhard Wilpert. *Safety Sci.* 49, 1–4. doi: 10.1016/j.ssci.2010.03.005

Cassano-Piche, A. L., Vicente, K. J., and Jamieson, G. A. (2009). A test of Rasmussen's risk management framework in the food safety domain: BSE in the UK. *Theor. Issues Ergon. Sci.* 10, 283–304. doi: 10.1080/14639220802059232

Curtis, R. (1995). *OA Guide to Outdoor Safety Management.* Available online at: https://www.princeton.edu/~oa/safety/safeman.html

Davenport, C. J. (2010). *Mangatepopo Coroners Report*. Auckland. Available online at: http://outdoorcouncil.asn.au/doc/Coroners_Report_OPC.pdf

Dekker, S. (2011). *Drift Into Failure: From Hunting Broken Components to Understanding Complex Systems*. Boca Raton, FL: CRC Press, Taylor & Francis Group.

Dekker, S., Cilliers, P., and Hofmeyr, J. H. (2011). The complexity of failure: implications of complexity theory for safety investigations. *Saf. Sci.* 49, 939–945. doi: 10.1016/j.ssci.2011.01.008

Department of Health (2006). *Safety First: A Report for Patients, Clinicians and Healthcare Managers.* Available online at: http://webarchive.nationalarchives. gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_ digitalassets/@dh/@en/documents/digitalasset/dh_064159.pdf

Drupsteen, L., Groeneweg, J., and Zwetsloot, G. (2013a). Critical steps in learning from incidents: using learning potential in the process from reporting an incident to accident prevention. *Int. J. Occup. Saf. Ergon.* 19, 63–77. doi: 10.1080/10803548.2013.11076966

Drupsteen, L., Groeneweg, J., and Zwetsloot, G. (2013b). Critical steps in learning from incidents: using learning potential in the process from reporting an incident to accident prevention. *Int. J. Occupat. Saf. Ergon.* 19, 63–77. doi: 10.1080/10803548.2013.11076966

Drupsteen, L., and Hasle, P. (2014). Why do organizations not learn from incidents? Bottlenecks, causes and conditions for a failure to effectively learn. *Accid. Anal. Prev.* 72, 351–358. doi: 10.1016/j.aap.2014.07.027

Goode, N., Finch, C., Cassell, E., Lenne, M. G., and Salmon, P. M. (2014a). What would you like? Identifying the required characteristics of an industry-wide incident reporting and learning system for the led outdoor activity sector. *Aus. J. Outdoor Educ.* 17, 2–15. Available online at: http://research.usc.edu.au/ vital/access/services/Download/usc:12724/SOURCE2 (Accessed December 16, 2016).

Goode, N., Salmon, P. M., Lenne, M., and Finch, C. F. (2014b). "A test of a systems theory-based incident coding taxonomy for risk managers," in *Advances in Safety Management and Human Factors*, Vol. 10, *Advances in Human Factors and Ergonomics 2014*, eds P. Arezes and P. Carvalho, 5098–5108.

Goode, N., Salmon, P. M., Lenne, M. G., Taylor, N. Z., and Finch, C. F. (2015). The UPLOADS Project: development of an Australian National Incident Dataset for Led Outdoor Activities. *Wilderness Environ. Med.* 26, 574–576. doi: 10.1016/j.wem.2015.04.006

Haddon, W. (1980). Advances in the epidemiology of injuries as a basis for public policy. *Public Health Rep.* 95, 411–421. Retrieved from: http://www.jstor.org/ stable/4596353

Hollnagel, E. (2004). *Barriers and Accident Prevention.* Aldershot: Ashgate.

Hollnagel, E. (2008). Risk + barriers = safety? *Saf. Sci.* 46, 221–229. doi: 10.1016/j.ssci.2007.06.028

Hollnagel, E. (2012). *FRAM: the Functional Resonance Analysis Method: Modelling Complex Socio-Technical Systems.* Aldershot: Ashgate Publishing, Ltd.

Jenkins, D. P., Salmon, P. M., Stanton, N. A., and Walker, G. H. (2010). A systemic approach to accident analysis: a case study of the Stockwell shooting. *Ergonomics* 53, 1–17. doi: 10.1080/00140130903311625

Jacobsson, A., Ek, A., and Akselsson, R. (2011). Method for evaluating learning from incidents using the idea of "level of learning". *J. Loss Prev. Process Industries* 24, 333–343. doi: 10.1016/j.jlp.2011.01.011

Jacobsson, A., Ek, A., and Akselsson, R. (2012). Learning from incidents – A method for assessing the effectiveness of the learning cycle. *J. Loss Prev. Process Industries* 25, 561–570. doi: 10.1016/j.jlp.2011. 12.013

Jacobsson, A., Sales, J., and Mushtaq, F. (2010). Underlying causes and level of learning from accidents reported to the MARS database. *J. Loss Prev. Process Industries* 23, 39–45. doi: 10.1016/j.jlp.2009.05.002

Johnson, C. W. (2003). *Failure in Safety-Critical Systems: A Handbook of Incident and Accident Reporting.* University of Glasgow.

Johnson, C. W., and Muniz de Almeida, I. (2008). Extending the borders of accident investigation: applying novel analysis techniques to the loss of the Brazilian space launch vehicle VLS-1 V03. *Saf. Sci.* 46, 38–53. doi: 10.1016/j.ssci.2006.05.007

Katsakiori, P., Sakellaropoulos, G., and Manatakis, E. (2009). Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Saf. Sci.* 47, 1007–1015. doi: 10.1016/j.ssci.2008.11.002

Kirwan, B. (2011). Incident reduction and risk migration. *Saf. Sci.* 49, 11–20. doi: 10.1016/j.ssci.2010.03.007

Le Coze, J. (2013). What have we learned about learning from accidents? Post-disasters reflections. *Saf. Sci.* 51, 441–453. doi: 10.1016/j.ssci.2012.07.007

Leveson, N. (2011). Applying systems thinking to analyze and learn from events. *Saf. Sci.* 49, 55–64. doi: 10.1016/j.ssci.2009.12.021

Lindberg, A. K., Hansson, S. O., and Rollenhagen, C. (2010). Learning from accidents – What more do we need to know? *Saf. Sci.* 48, 714–721. doi: 10.1016/j.ssci.2010.02.004

Lundberg, J., Rollenhagen, C., and Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals. *Saf. Sci.* 47, 1297–1311. doi: 10.1016/j.ssci.2009.01.004

Lundberg, J., Rollenhagen, C., and Hollnagel, E. (2010). What you find is not always what you fix—How other aspects than causes of accidents decide recommendations for remedial actions. *Accid. Anal. Prev.* 42, 2132–2139. doi: 10.1016/j.aap.2010.07.003

Lundberg, J., Rollenhagen, C., Hollnagel, E., and Rankin, A. (2012). Strategies for dealing with resistance to recommendations from accident investigations. *Accid. Anal. Prev.* 45, 455–467. doi: 10.1016/j.aap.2011.08.014

Newnam, S., and Goode, N. (2015). Do not blame the driver: a systems analysis of the causes of road freight crashes. *Accid. Anal. Prev.* 76, 141–151. doi: 10.1016/j.aap.2015.01.016

Nielsen, K. J., Carstensen, O., and Rasmussen, K. (2006). The prevention of occupational injuries in two industrial plants using an incident reporting scheme. *J. Saf. Res.* 37, 479–486. doi: 10.1016/j.jsr.2006.06.005

Pidgeon, N., and O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. *Saf. Sci.* 34, 15–30. doi: 10.1016/S0925-7535(00)00004-7

Pless, B. (2008). Surveillance alone is not the answer. *Inj. Prev.* 14, 220–222. doi: 10.1136/ip.2008.019273

Ramanujam, R., and Goodman, P. S. (2011). The challenge of collective learning from event analysis. *Saf. Sci.* 49, 83–89. doi: 10.1016/j.ssci.2010.03.019

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Saf. Sci.* 27, 183–213. doi: 10.1016/S0925-7535(97)00052-0

Read, G. J. M., Salmon, P. M., Lenné, M. G., and Jenkins, D. P. (2015). Designing a ticket to ride with the Cognitive Work Analysis Design Toolkit. *Ergonomics* 58, 1266–1286. doi: 10.1080/00140139.2015.1013576

Major Accidents Reporting System (2012). Available online at: https://emars.jrc.ec.europa.eu/

Aviation Safety Reporting System (2015). *ASRS Program Briefing.* Available online at: http://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing2015.pdf

Rollenhagen, C. (2011). Event investigations at nuclear power plants in Sweden: reflections about a method and some associated practices. *Saf. Sci.* 49, 21–26. doi: 10.1016/j.ssci.2009.12.012

Rollenhagen, C., Westerlund, J., Lundberg, J., and Hollnagel, E. (2010). The context and habits of accident investigation practices: a study of 108 Swedish investigators. *Saf. Sci.* 48, 859–867. doi: 10.1016/j.ssci.2010.04.001

Salmon, P. M., Cornelissen, M., and Trotter, M. J. (2012). Systems-based accident analysis methods: a comparison of Accimap, HFACS, and STAMP. *Saf. Sci.* 50, 1158–1170. doi: 10.1016/j.ssci.2011.11.009

Salmon, P. M., Goode, N., Archer, F., Spencer, C., McArdle, D., and McClure, R. J. (2014a). A systems approach to examining disaster response: using Accimap to describe the factors influencing bushfire response. *Saf. Sci.* 70, 114–122. doi: 10.1016/j.ssci.2014.05.003

Salmon, P. M., Goode, N., Lenn,é, M. G., Finch, C. F., and Cassell, E. (2014b). Injury causation in the great outdoors: a systems analysis of led outdoor activity injury incidents. *Accid. Anal. Prev.* 63, 111–120. doi: 10.1016/j.aap.2013.10.019

Salmon, P. M., Goode, N., Taylor, N. Z., Lenne, M. G., Dallat, C., and Finch, C. F. (2016a). Rasmussen's legacy in the great outdoors: a new incident reporting and learning system for led outdoor activities. *Appl. Ergon.* 59, 637–648. doi: 10.1016/j.apergo.2015.07.017

Salmon, P. M., Read, G., Stanton, N. A., and Lenné, M. (2013). The Crash at Kerang: investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings. *Accid. Anal. Prev.* 50, 1278–1288. doi: 10.1016/j.aap.2012.09.029

Salmon, P. M., Walker, G. H., Read, G. J. M., Goode, N., and Stanton, N. A. (2016b). Fitting methods to paradigms: are ergonomics methods fit for systems thinking? *Ergonomics* 22, 1–12. doi: 10.1080/00140139.2015.1103385

Salmon, P. M., Williamson, A., Lenne, M., Mitsopoulos-Rubens, E., and Rudin-Brown, C. M. (2010). Systems-based accident analysis in the led outdoor

activity domain: application and evaluation of a risk management framework. *Ergonomics* 53, 927–939. doi: 10.1080/00140139.2010.489966

Stanton, N. A. (2016). On the reliability and validity of, and training in, ergonomics methods: a challenge revisited. *Theor. Issues Ergon. Sci.* 17, 345–353. doi: 10.1080/1463922X.2015.1117688

Stanton, N. A., and Stevenage, S. V. (1998). Learning to predict human error: issues of acceptability, reliability and validity. *Ergonomics* 41, 1737–1756. doi: 10.1080/001401398186162

Svedung, I., and Rasmussen, J. (2002). Graphic representation of accidentscenarios: mapping system structure and the causation of accidents. *Saf. Sci.* 40, 397–417. doi: 10.1016/S0925-7535(00)00036-9

Taylor, N. Z., Goode, N., Salmon,., P. M., Lenne,., M. G., and Finch, C. F. (2015a). "Inter-rater reliability of a causal factor taxonomy that uses systems theory designed for the outdoor domain," in *Paper Presented at the AHFE2015, 1st International Conference on Human Factors in Sports and Outdoor Recreation* (Las Vegas, NV).

Taylor, N. Z., Goode, N., Salmon, P. M., Lenne, M. G., and Finch, C. F. (2015b). "Which code is it? Inter-rater reliability of systems theory-based causal factor taxonomy for the outdoor sector," in *Paper Presented at the 19th Triennial Congress of the International Ergonomics Association* (Melbourne, VIC).

Underwood, P., and Waterson, P. (2014). Systems thinking, the Swiss Cheese Model and accident analysis: a comparative systemic analysis of the Grayrigg train derailment using the ATSB, AcciMap and STAMP models. *Accid. Anal. Prev.* 68, 75–94. doi: 10.1016/j.aap.2013.07.027

van Mulken, M., Clacy, A., Grant, E., Goode, N., Finch, C. F., Stevens, E., et al. (2016). *The UPLOADS National Incident Dataset The First Twelve Months: 1st June 2014 to 31st May 2015.* Maroochydore: Centre for Human Factors and Sociotechnical Systems. https://uploadsproject.files.wordpress.com/2016/03/first-12-month-report_final.pdf (Accessed December 15, 2016).

Vastveit, K. R., Boin, A., and Njå, O. (2015). Learning from incidents: practices at a Scandinavian refinery. *Saf. Sci.* 79, 80–87. doi: 10.1016/j.ssci.2015.05.001

Vicente, K. J. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work.* CRC Press.

Vincent, C. (2004). Analysis of clinical incidents: a window on the system not a search for root causes. *Qual. Saf. Health Care* 13, 242–243. doi: 10.1136/qshc.2004.010454

Check for
updates

# Technology as Teammate: Examining the Role of External Cognition in Support of Team Cognitive Processes

Stephen M. Fiore[1]* and Travis J. Wiltshire[1,2]

[1] Department of Philosophy and Institute for Simulation & Training, University of Central Florida, Orlando, FL, USA,
[2] Dynamical Systems Lab, Department of Psychology, University of Utah, Salt Lake City, UT, USA

In this paper we advance team theory by describing how cognition occurs across the distribution of members and the artifacts and technology that support their efforts. We draw from complementary theorizing coming out of cognitive engineering and cognitive science that views forms of cognition as external and extended and integrate this with theorizing on macrocognition in teams. Two frameworks are described that provide the groundwork for advancing theory and aid in the development of more precise measures for understanding team cognition via focus on artifacts and the technologies supporting their development and use. This includes distinctions between teamwork and taskwork and the notion of general and specific competencies from the organizational sciences along with the concepts of offloading and scaffolding from the cognitive sciences. This paper contributes to the team cognition literature along multiple lines. First, it aids theory development by synthesizing a broad set of perspectives on the varied forms of cognition emerging in complex collaborative contexts. Second, it supports research by providing diagnostic guidelines to study how artifacts are related to team cognition. Finally, it supports information systems designers by more precisely describing how to conceptualize team-supporting technology and artifacts. As such, it provides a means to more richly understand process and performance as it occurs within sociotechnical systems. Our overarching objective is to show how team cognition can both be more clearly conceptualized and more precisely measured by integrating theory from cognitive engineering and the cognitive and organizational sciences.

Keywords: team cognition, macrocognition in teams, external team cognition, teamwork, taskwork, offloading, scaffolding

## INTRODUCTION

Organizations are often characterized as complex sociotechnical systems that require effective coordinative and collaborative cognitive processes across individuals and teams in order to meet their goals. As such, research in team cognition has become increasingly prevalent over the past decade. Team cognition is a broad area of research meant to explore the manifestation of cognition in the context of teamwork (Salas and Fiore, 2004; Letsky et al., 2008; Salas et al., 2012; Turner et al., 2014). This includes understanding how memory influences teams (e.g., transactive memory

systems, Lewis and Herndon, 2011), or how cognitive constructs, such as mental models, can provide explanatory value to inform team functioning (e.g., shared mental models, DeChurch and Mesmer-Magnus, 2010b). Other processes such as attention and decision making, as they arise in teams, have also been studied [e.g., distributed cognition, Hutchins, 1995a; distributed situation awareness (DSA), Stanton, 2016]. A significant amount of research on these topics has been able to inform our understanding of teams and how, for example, training (Cooke and Fiore, 2009) or system design (Kiekel and Cooke, 2004; Bowers et al., 2006) can be improved.

We suggest, however, that what constitutes cognition in the organizational sciences is too often narrowly construed. This potentially leads to an incomplete understanding of team processes and the many factors leading to successful performance, particularly when teams are made up of a hybrid of humans and technology. Specifically, despite a large body of research, there is less attention paid to external cognition, that is, artifacts or material objects used in service of team cognition, or technologies supporting their development and use, and how these relate to team effectiveness. In the more general study of teams, there have been discussions of teams and their relation to technology. For example, when viewing teams as a human-technology system (Kozlowski et al., 2015), researchers describe how the technological sub-system is an important component to understanding the kinds of emergent processes typically related to team effectiveness (e.g., cohesion or collective efficacy). Others have noted how the technology, itself, can shape communicative and coordinative interactions and, thus, substantially influence team process (Bell and Kozlowski, 2012). Nonetheless, studies of technology, and the artifacts it helps teams produce, is under represented in the team cognition literature.

As evidence of this lack of inquiry within the field of team cognition, recent reviews have not made mention of artifacts or associated terms, or even of technology, in any substantial way. For example, although drawing from multiple disciplines and providing what is described as a "cross-domain review" on the measurement of team cognition, there was no mention of how external cognition factors like material objects, artifacts, or technology should be considered as part of the team process (Wildman et al., 2014). In a review on the role of team knowledge in understanding collaborative processes, despite a comprehensive coverage of the ways knowledge is conceptualized, there is no mention of how these external cognition factors artifacts relate to knowledge construction and use, nor how they should fit within team cognition research (Wildman et al., 2012). Similarly, in a meta-analysis of team cognition constructs, these factors were not considered in any of the classifications that examined the relationship between team cognition and performance outcomes (Turner et al., 2014).

What is striking about these and earlier similar articles (e.g., DeChurch and Mesmer-Magnus, 2010a,b), is that many of the studies making up the foundation for these reviews, in some form or another, used technologies that would create or need artifacts for task completion. As an example, this could include project management type tasks where planning required the creation or use of artifacts, or computer-based experimental tasks (e.g.,

simulations of aviation necessitating use of diagrams), to even just technologies supporting information sharing and storage (e.g., chat boards). Our point is that there is tremendous potential in considering these externalized cognition factors as a relevant element of team cognition. From this, research can examine the degree to which it may moderate or mediate any number of team process and performance outcomes and help us understand and improve team cognition.

In short, we suggest that team cognition research lacks the conceptual scaffolds necessary to examine how artifacts and associated technologies are related to team process and performance. To redress this gap, we integrate a set of constructs under the general label of external cognition to describe how the concept of artifacts, and the technology supporting their development and use, have been discussed as a foundational part of collaboration across a number of fields. With that as a stepping off point, we then show how distinctions between teamwork and taskwork, arising from organizational theory on team training, and differences between offloading and scaffolding cognition, arising from the cognitive sciences, can be united to provide a framework that advances team cognition research. Our goal is to show how these provide explanatory value to team cognition theory by helping to conceptualize technology as teammate.

This paper consists of two major sections, each with two subsections. First we provide an overview of the general idea of technology in team cognition in the context of research and theory on complex collaborative environments where technology is inherent and cognition is often externalized. Second, using the general label of "artifact" we describe how external forms of cognition have been examined in a variety of settings so as to provide evidence for the reach of this idea and how it has been related to cognition and collaboration. This initial half of the paper provides the foundational literature on which we build the argument for examining technology, broadly construed, as part of a team. The latter half works to integrate ideas from organizational research on teams, and concepts from cognitive science, to provide a novel means through which to understand team cognition. Specifically, in the third section, we discuss the distinction between "teamwork" and "taskwork" – ideas that have yet to be integrated with the external cognition perspective. Fourth, we bring in ideas from cognitive science about offloading and scaffolding cognition to show how these help us more finely distinguish between forms of external cognition in the context of teams. Within these sections we provide guidelines and research questions devised around technology in support of external cognition to help researchers examine teams as socio-technical systems.

## COGNITION, TECHNOLOGY, AND TEAMS

In an age of ubiquitous technology, the study of team cognition needs research that more closely examines our assumptions about what is cognition and its manifestation through, and within, technology in the modern workplace. This is necessary to develop the next phase of team cognition research for the organizational

sciences. Indeed, there have been recent calls for research on teams to improve understanding human-system issues arising from the team-technology integration. For example, Bell and Kozlowski (2012) called out the lack of studies in organizational research that have fully examined the complementarity between technology and team interaction and how they lead to emergent states. In this context, they specifically labeled such issues as one of the important themes for future research on teams. More recently, Kozlowski et al. (2015) noted the criticality of understanding how workflow within teams, interacts with technology to influence cognition and behavior. They highlight the need for more research on team design and, included in this, is a need for research that examines how technologies can help or hinder numerous factors related to team cognitive factors (e.g., information sharing and distribution).

Toward this end, drawing from research focusing on the intersection of cognition and technology as it occurs in naturalistic and dynamic organizational contexts (Cacciabue and Hollnagel, 1995; Pennathur et al., 2008; Jenkins et al., 2011; Fiore, 2012; Cooke et al., 2013; Lee and Kirlik, 2013; Gorman, 2014), we integrate theory from cognitive engineering with the cognitive and organizational sciences in order to help team researchers more fully conceptualize cognition in its varied forms. We show how the next phase of team cognition research can be pursued as a form of team-technology hybrid wherein we can come to better understand the tight coupling between the individual, the team, and the technologies they rely upon.

Our main argument is that understanding team cognition as it occurs in real-world work settings requires an expanded view where cognition is seen as distributed and context dependent in a social environment in which artifacts often support cognitive functions (Suchman, 1987, 2007; Hutchins, 1995a; Clancey, 1997; Hollnagel, 2002). Specifically, we advance the notion that artifacts support cognition by enabling the transition and development of internalized knowledge held by team members to externalized knowledge held at the team-level (Fiore et al., 2010b; Rentsch et al., 2010, 2014). We draw from a diverse body of research and theory to emphasize that the functions of cognition can, and must be, viewed as sometimes occurring, not just "in" the head, but also "outside the head"; that is, viewing cognition in a broader context as distributed across the boundaries of brains, bodies, and environment (Fiore, 2012; Cooke et al., 2013; Gorman, 2014). We describe DSA theory (e.g., Stanton, 2016), interactive team cognition (ITC) theory (e.g., Cooke and Gorman, 2009; Cooke et al., 2013), and macrocognition in teams (MiTs) theory (Fiore et al., 2008, 2010b,c) from cognitive engineering, and extended cognition theory from cognitive science (Clark and Chalmers, 1998; Clark, 2001a,b), to better understand the increasingly prevalent role technology plays as a form of external cognition in complex collaborative work domains.

The combination of these perspectives provides a strong foundation from which the organizational sciences can begin to consider and measure external team cognition in order to contribute to team theory and practice and, in turn, increase organizational effectiveness. We now turn to a discussion of theory that has broadly considered how contextual factors, like technology, play a role in team process.

## Considerations of Context and Team Cognition

The 20th century saw tremendous gains in organizational productivity thanks to numerous technological advances. As mechanization began to dominate in the early decades, work practices changed and humans adapted to these new systems. Importantly, organizational scientists studying these changes recognized that not all adaptations were equal. In the middle part of the century, researchers with the Tavistock Institute observed innovative work practices that moved beyond bureaucratization and mechanization to create a new form of work. In the British mining industry, where technology had made tremendous inroads, some workers had developed a higher form of collaboration between themselves and their technology (for a discussion, see Trist, 1981). Viewed as a sub-system of the organization within which autonomy had been enhanced, it could lead to greater group cohesion, self-regulation and coordination as teams developed new practices for working with each other and the new technologies. This was seen as an important alternative to Tayloresque and Weberian approaches in that, for organizational design, "the best match would be sought between the requirements of the social and technical systems" (p. 9). In many respects, this revolutionized organizational theory by introducing systems thinking into the lexicon and helping to produce a more holistic view of the interactions between, people, machines, and the environmental context in which they operate (Trist, 1981).

We open this section with this brief historical perspective because, although socio-technical systems theory was an important part of organizational research, and originated from a study of groups working with technology, this perspective had less influence on the study of teams. Research in teams throughout most of the 20th century focused more on the social than the technical (e.g., Guzzo and Dickson, 1996). Furthermore, with the advent of the cognitive revolution in the organizational sciences, we saw an infusion of research on the interaction of the social and the cognitive (Hinsz et al., 1988, 1997; Lord and Maher, 1991; Larson and Christensen, 1993), but, still, with little incorporation of technology's role in teams. Rather, this led to the emergence of the study of team cognition and the manifestation of cognition within and across individuals during complex and dynamic interactions (e.g., Salas and Fiore, 2004).

From this we gained significant understanding of how social and cognitive factors influence process and performance. For example, a tremendous amount of research has studied the relationship between team knowledge, such as shared mental models, and team outcomes (DeChurch and Mesmer-Magnus, 2010b). Research has also studied how coordination is altered by expertise within the team (e.g., Faraj and Sproull, 2000; Espinosa et al., 2004, 2007), or how coordinative mechanisms are necessary for reaching shared goals or achieving desired performance outcomes (Gittell and Weiss, 2004; Gittell, 2006; Brodbeck et al., 2007; Okhuysen and Bechky, 2009). In brief, there has been a pervasive emphasis on the role of stable mental constructs such as shared knowledge and/or coordination processes. But these cognitive structures are still abstract, subjective, internal,

and provide a restricting view of cognition to the organizational sciences (e.g., Mitchell et al., 2011). This research now transcends disciplines and many theories, methods, and domains are part of team cognition research (Salas et al., 2012).

Despite these theoretical and empirical advances in numerous areas, most organizational research on teams has not taken into account how the environment in general, and technology, in particular, interacts with individual and team cognition. Recent efforts have called for stronger integration of these approaches (e.g., Rico et al., 2011) as well as for a broader perspective on what is meant by team cognition and how interaction dynamics and context are related to team effectiveness (Fiore et al., 2010a; Cooke et al., 2013; Cooke, 2015). Along these lines, we argue that team research in the organizational sciences will benefit from theories emerging in other disciplines that more fully account for the role of contextual factors in team cognition in general, and the role of technology, in particular. Generally, these theories consider cognition as something more than that which goes on "inside the head"; rather, cognition is something that can be studied both *inside and outside the head* as team members interact with each other and their technology. We next briefly review some of this theorizing.

## Context and Behavior When Interacting with Technology

In early theorizing in this area, research on *situated cognition*, by social anthropologist, Suchman (1987), argued that the agent and the environment have to be included in theorizing about cognition. Her research emphasized the role of context in cognition and the use of ethnomethodology to analyze human activity arising between a person and the setting in which that activity takes place. From this, researchers began to recognize relational coupling between situation and action, where meaning is constructed within particular contexts (Fiore, 2013).

Even information processing theorists made a claim for the value of understanding cognition as situated (Vera and Simon, 1993). They argued that symbolic and representational approaches could explain interactions with complex work systems. From this perspective, simulation of cognitive activity can be conceived of as occurring within and across individuals and the representational systems on which they rely (see also Larkin and Simon, 1987).

Coming out of research on cognitive engineering, DSA theory was another to examine context to place emphasis on understanding the social-technical system in its entirety as the unit of analysis (Stanton et al., 2006; Stanton, 2016). While DSA proposes that researchers delineate between their adopted unit of analysis such as 'in mind,' 'in world,' or 'in-interaction,' the focus of DSA is typically on the behavioral interactions that facilitate the transaction of awareness amongst agents in a socio-technical system, whether those are social sub-systems (e.g., individual humans and teams) or technical sub-systems (e.g., technologies, interfaces, artifacts, displays, etc.). In this case, situation awareness refers to holding information regarding the status of a given situation. But DSA differs from traditional notions of SA (Endsley, 1995) in that is does not assume that SA can be held only "in-mind" of humans, but rather it can be

distributed across the technologies as well and is available to the human as needed.

When considering SA in teams, empirical work comparing a DSA approach to a traditional team cognition approach on shared SA, found that teams who had awareness that was more differentially distributed across team members (as shown by concept maps) performed better than teams who shared more information and held largely the same awareness on a rogue vehicle detection task (Kitchin and Baber, 2016). In another example, teams working in anesthesia management were shown to explicitly rely on their interactions with artifacts such as computer monitors and whiteboards, as well as their teammates to gain the appropriate awareness that allowed them to perform their duties effectively (Fioratou et al., 2016). Other studies have similarly shown that it is more important that awareness is distributed across team members and their technologies and that such cases often exhibit improved task performance (e.g., Bourbousson et al., 2011; Sorensen and Stanton, 2013).

Coming out of the cognitive sciences, others have similarly conceptualized and examined team cognition and behaviors at the collective level. Specifically, ITC theory (Cooke and Gorman, 2009; Cooke et al., 2013; Cooke, 2015) draws from post-information processing perspectives of individual cognition, such as embodied cognition and activity theory. ITC views team cognition more dynamically, as an activity engaged by teams over time and, in line with earlier views of situated cognition (e.g., Suchman, 1987), sees cognition as inseparable from context. Similar to DSA, an important tenant of ITC is that team cognition needs to be examined at the level of the team (e.g., communication; Cooke et al., 2008, 2004). Finally, it differs primarily from traditional theories of team cognition by arguing that performance differences can be more accurately understood, not by knowledge differences in the team (e.g., shared mental models), but in the behavioral interactions (Cooke et al., 2009; Gorman et al., 2010).

Empirical evidence for ITC theory comes from findings where the disruption of interactions patterns during task training actually improve later performance when compared to those whose interaction patterns were not disrupted (Gorman et al., 2010). Teams that were disrupted learned to adapt interaction behaviors that later proved beneficial. Other results show that, while team performance increases across a full series of performance events, changes to team knowledge occur primarily during earlier events, whereas, changes and refinements to the team's interactive processes occurs during more of the missions (Cooke et al., 2001). This suggests that the collective and interactive behaviors are what is driving the continued team performance improvements, rather than the continued development of task knowledge.

In sum, the argument that theorizing on collaborative cognition should account for contextual and technological factors, has been an important part of research on teams operating in complex settings. These views converge on the perspective that cognition can occur at the intersection of the individual, the team, their technology, and the environment, to influence their behaviors in context. This work makes strides in helping us see how features and components of tasks can be

distributed across team member's internal cognitive systems, the collective external cognitive system of the team, as well as across artifacts and technologies in the environments in which they interact (Zhang and Norman, 1994; Zhang, 1998; Hutchins, 1999; Stanton et al., 2006; Clark, 2008; Fiore et al., 2010b; Cooke et al., 2013).

We build from this to argue that external cognition as part of that context, whether it be physical, mechanical, technological or otherwise, needs to be recognized and measured as a part of team cognition. This, then, can be used to help us understand and measure *where* the team is being supported by these as well as *how*. In this way, we add to team cognition research by focusing on the ways in which teams collaborate with each other *and* with/through technology. We next discuss how MiTs theory, an approach aligned with these perspectives, can advance research on teams. We focus on the role of technology in support of external cognition to provide theoretical guidance that can facilitate empirical work in this area.

## External Cognition and Macrocognition in Teams

Researchers studying cognition embedded in rich, real-world environments, developed the concept of *macrocognition,* a term that embodies a shift away from the traditional micro-view of cognition to describe how cognition operates when faced with complexity (Hollnagel, 2002). Broadly, macrocognition includes the ideas that: (a) across natural and artificial cognitive systems, the process and product of cognition will be distributed; (b) cognition is not self-contained and finite, but a continuance of activity; (c) cognition is contextually embedded within a social environment; (d) cognitive activity is not stagnant, but dynamic; and (e) artifacts aid in nearly every cognitive action (Hollnagel, 2002; Klein et al., 2003, 2006; Fiore, 2012). These ideas provide important additional explanatory power by providing an enhanced appreciation of how interaction unfolds in dynamic and contextually rich settings.

Macrocognition in teams theory is an interdisciplinary integration of much of this prior research on collaborative cognition that emphasizes both internalized and externalized cognition and the role of artifacts in collaboration (Fiore et al., 2008, 2010b,c). In addition to considering how, for example, shared memory structures support teamwork (e.g., understanding how team mental models help sequence actions), MiTs theory focuses on ways in which internalized knowledge is transformed to externalized knowledge by both individual and team-level cognitive processes for the purposes of knowledge coordination (Fiore et al., 2010b). In this way, it addresses how teams externalize cognition to collaboratively build knowledge through the transformation of data to information to knowledge in service of team problem solving (Fiore et al., 2010b). The macrocognitive view is particularly relevant to this paper given that prior theorizing specifically emphasized how individuals and teams deal with complexity via reliance on technology (e.g., Hollnagel, 2002; Klein et al., 2006).

Foundational to MiT theory is the notion of *extended cognition* (Clark and Chalmers, 1998; Clark, 2001a). Similar to the theorizing discussed earlier, this perspective argues that the brain is inextricably coupled to one's external environment and often relies on this coupling for many complex tasks. The extended cognition perspective also posits that some of what is normally construed as cognition localized "within the head," can also occur beyond the boundaries of the head, that is, as externalized cognition. Two simple examples of extended cognition include note-taking during a lecture and working out a mathematical problem on paper. Broadly, the former is an act of "remembering" in the sense that this is a type of external storage to which one can later refer. The latter is an act of "cognition" in the sense that the mental effort required to solve the problem is off-loaded onto the environment (i.e., calculations are not all done entirely mentally).

More generally, if a given task requires the temporarily formed and synergistic coalition of the body's sensorimotor systems and neural circuits, as well as artifacts and/or other people in the environment, then it is difficult to relegate the functions of cognition to just occurring within the head (Anderson et al., 2012). Note that the extended view of cognition does not claim that the brain is not playing a crucial role in cognition. Rather, the point is that the role of the brain, at least in this respect, is to act "as a mediating factor in a variety of complex and iterated processes which continually loop between brain, body, and technological environment" (Clark, 2002, p. 24). Through this theoretical lens, cognitive functions can be construed of as extending outside of the body, that is, externalized. Of course, to include artifacts as part of cognition is contingent upon the notion that their use must be available when needed and accessed in ways analogous to traditional retrieval mechanisms (Clark, 2001b). This sociotechnical system is the foundation from which solutions to complex problems can emerge (Fiore et al., 2008).

In their theorizing on MiTs, Fiore et al. (2010b,c) wove this into an elaboration of the functional role externalized cognition plays in collaborative problem solving. Motivation for this claim stems from the notion that "the degree the team-task requires the construction of a shared understanding, external representational tools can act as a scaffolding to facilitate the building of that shared representation" (Fiore and Schooler, 2004, p. 134). Building on this, in MiT theory, externalized cognition can be a focal point for team discussion and elaboration, and can support analysis of ideas put forth, and potential solutions, by helping members attend to key details articulated in the externalization. In this view, externalized cognition is particularly useful when teams are supported by technology; that is, by sociotechnical systems devised to help members deal with the tremendous variety of data and information with which they are confronted when dealing with complex problems (cf. Klein et al., 2003, 2006).

A key gap in the theorizing on MiT theory, though, is that it does not fully articulate the richness of what is meant by externalized cognition. Although it describes external cognition as an important component of knowledge building in teams, the specific ways in which external cognition can manifest itself, and how it plays a role in extending team cognition, need to be better articulated. We next address this gap in MiT theory via explication of artifacts as externalized cognition and articulation of the specific ways these play a role in different aspects of team cognition. Toward this end, we summarize some of the prior research on which the MiT theory was built and which specifically

focuses on the idea of externalized cognition and role artifacts play when collaborating.

## Artifacts and Technological Support as Externalized Team Cognition

Although we have claimed the essentiality for examining artifacts as a part of team cognition, we so far, have yet to elaborate on what we mean when we refer to artifacts and the evidence for their value to team cognition. Therefore, in this section, we provide a foundation for conceptualizing artifacts, and the varied ways in which they've been viewed, as a form of external cognition. In addition, we also review various technologies that have been developed to support teams in a number of domains that are characteristic of artifacts that facilitate external cognition. This section illustrates how evidence for this area of inquiry has been independently developing in a variety of fields that do not always influence each other and show how to leverage these developments to integrate ideas on external cognition with team research in the organizational sciences.

The notion of the *cognitive artifact* emerged in studies of design and human–computer interaction and was characterized by Norman (1991, p. 17) as an "artificial device designed to maintain, display, or operate upon information in order to serve a representational function." Importantly, there is a long history in the social sciences of conceptualizing artifacts as a means for supporting human capabilities. As noted by Norman (1991), a number of theoretical positions emerging in the 20th century, such as "activity theory" or "situated action," focused on the role of the natural and artificial environment in enhancing human abilities (for a review, see Fiore, 2013). Even early work on information processing theory discussed how representations as external symbol systems foster complex cognitive processes (Larkin and Simon, 1987). Here, diagrammatic representations were said to group related information and minimize problem space search and also support perceptual inferences about processes. These and related features of externalization were argued to make cognitive processes more computationally efficient. This early thinking influenced research where the focus was on cognition in work contexts as well as in learning and training research. In these varied settings, the concept of an artifact has fallen under a number of labels, but all related in the sense that they are a form of external cognition. We next briefly review these in turn.

### Artifacts in Distributed Cognition

An influential early theory with representational information at its core is Hutchins (1995a,b) theorizing on distributed cognition. The primary argument is that cognitive processes are not only internal, but are also spread across task and environmental artifacts, as well as team members. Heavily based on information processing theory, cognitive processes were said to act on these representations via computation of some form to transform understanding. Cognitive artifacts were defined as "physical objects made by humans for the purpose of aiding, enhancing, or improving cognition" (Hutchins, 1999, p. 126). With this, the focus was on the interaction of distributed structures in a broader cognitive system. As one example,

Hutchins used cockpit technology (e.g., attitude indicator) and the aviation crew, to describe a distributed cognitive system. In these specific settings, some have even discussed the idea that automation technology be construed of as a teammate (Hoeft et al., 2006). This expanded the boundaries of how cognition can be analyzed – with distribution encompassing processes across time, as well as across the team, and internal and external cognitive structures in humans, and their supporting technology.

This was further detailed in the context of human–computer interaction research, where an ethnographic approach was used to study the use of digital artifacts that trace histories of interaction (Hollan et al., 2000). Here, distributed cognition was examining the interplay of internalized and externalized cognition "involving coordination at many different time scales between internal resources—memory, attention, executive function—and external resources—the objects, artifacts, and at-hand materials constantly surrounding us" (Hollan et al., 2000, p. 177).

Research in healthcare teams also examined the role of cognitive artifacts in supporting coordination across team members (Nemeth et al., 2004, 2006; Rambusch et al., 2004). In this context, it was shown how technologies assisted teams in the form of externalizations such as team schedules, lists, display boards, and patient records. Similarly, in the context of emergency rooms, externalization of cognition, through the use of whiteboards, was shown to support coordinating responsibilities and resources. In short, artifacts in the form of visual representations, act as aids to memory and provide information directly perceivable by members of the team to facilitate collaboration. These forms of external cognition helped teams maintain a shared overview of the total team activity distributed across time, location, and across different technologies (Nemeth et al., 2004). External cognition has also been shown to help teams dynamically make decisions and identify potential problems that might arise in their task (Xiao et al., 2007). This and related work has been used to help system designers understand how artifacts could be transitioned to digitally based forms to create a more resilient system.

Collectively, this works shows how artifacts support activities like team planning by mediating collective work and the management of resources (Nemeth et al., 2004). They further elucidate how this can vary as a function of *who* was using a given artifact and *where* (Rambusch et al., 2004). Taken together, this research provides a foundation for seeing teams, their technology, and the resultant externalizations, as a distributed cognitive system (cf. Hutchins, 1995b).

### Boundary Objects in Organizational Research and Computer Supported Cooperative Work (CSCW)

In the organizational sciences, the concept of materiality and sociomaterialy are often used to capture how some artifact, loosely defined, influences, and is influenced by, work processes. This body of research examines organizational functions at a broad level (e.g., finance), and how technology relates to that (e.g., spreadsheet software). And much work has gone in discussion and debate on what is meant by materiality and related terms (Leonardi, 2010, 2012). Theoreticians have debated how to

conceptualize this idea and its relation to the material part of organizations. Here they argue that, "whereas materiality might be a property of a technology, sociomateriality represents that enactment of a particular set of activities that meld materiality with institutions, norms, discourses, and all other phenomena we typically define as 'social'" (Leonardi, 2012, p. 34).

Relevant to this paper, reviews in the organizational sciences note that most studies in areas relevant to cognition (e.g., decision-making, strategic thinking), *have not* considered how technology influences these complex processes (Orlikowski and Scott, 2008). Similarly, some have argued that organizational research needs to better integrate ideas about how information technology and the materiality it affords, is related to the functions and processes of organizations (Leonardi and Barley, 2008). This work shows the far reaching recognition that externalizations provide a powerful means of connecting people.

Despite the conceptual connection of such ideas to the notion of artifacts, socio-materiality operates at a level above teamwork. That is, it transcends work in teams and represents objects that connect, not necessarily individuals within a team, but groups of people within an organization, and even entire communities of practice. As such, this body of research has not had an influence on, let alone been integrated with, team cognition. But fields that focus more on technology and its relation to team functions [e.g., Information Systems, Computer Supported Cooperative Work (CSCW)], come close to addressing this gap through the development of the concept of *boundary objects*. As such, we next turn to a description of research in boundary objects to set the stage for how this can be related to team cognition.

Within the field of CSCW, a significant amount of research has been on the development and use of what were termed material resources (see Blomberg and Karasti, 2013, for a review). These ranged from artifacts as simple as paper documents to computer displays and whiteboards and maps. Early work in this area showed how these help collaborators align their activities by drawing attention to coordination needs (Suchman and Trigg, 1991; Heath and Luff, 1992). Some have used the generic label of "shared representation" to capture this concept. These are external representations that arise in collaborations and can vary in meaning and relevance depending on context in which they are used (see de Vries and Masclet, 2013). Out of such work arose a particular form of sociomaterialy, known as boundary objects. Originating in research on scientific work, these were described as practical artifacts that mediate interaction across diverse groups and communities of practice with varying expertise and perspectives (Carlile, 2002, 2004; Yakura, 2002; Hecker, 2012). These "tangible" artifacts were shown to act as a bridge from which communication and coordination occur, thus facilitating not only the transfer of knowledge from an individual to a team level, but also the maintenance of shared representations (Yakura, 2002; Nicolini et al., 2012; Stigliani and Ravasi, 2012).

This concept has had an influence in a number of domains as it has been adopted by researchers in the organizational and information sciences (see Lee, 2007 for a review). Early research on boundary objects suggests that they foster cooperation between diverse communities of stakeholders through creation of a shared identity (Star and Griesemer, 1989; Star, 2010).

Additionally, boundary objects were seen as a means of both knowledge transfer, and a method for translating meaning across an organization utilizing shared information systems (Carlile, 2004). Some have looked at this in the context of, not just the development of information systems, but also their implementation (Doolin and McLeod, 2012). Here, it was argued that communities of practice needed to develop competencies *about* boundary objects so that those working in these settings could make them useful (Levina and Vaast, 2005).

From this, CSCW researchers described the development of "common information spaces" that help make explicit "the interrelationships between information, workers, and artifacts. . . [and] involve the joint interpretation of and the meaning attributed to these artifacts and representations" (Blomberg and Karasti, 2013, p. 382). And out of this came the notion of "coordinative artifacts" that were seen as essential to collaboration in complex cognitive work (Schmidt and Wagner, 2004; Lee, 2007). These were argued to reduce the amount of articulation of what needed to be done by specifying division of labor as well as sequencing/ordering of activities (Bardram and Bossen, 2005). Along with this was the need for active negotiation about a boundary object in order to develop shared understanding (Lee, 2007). These served different roles dependent upon the task needs – ranging from the simple, such as including ideas or compiling ideas (e.g., tables), to the more complex, such as structuring ideas (e.g., concept maps). These were said to serve either a syntactic function to help collaborators transfer knowledge via a common vocabulary, or a semantic function that helps identify differences in knowledge to create shared knowledge (Carlile, 2004). In brief, these allow collaborators to "record, organize, explore and share ideas; introduce concepts and techniques; create alliances; create a venue for the exchange of information; augment brokering activities; and create shared understanding about specific problems" (Jirotka et al., 2013, p. 668).

Research has also examined how these forms of external cognition help orient team members in complex decision making tasks. In a study of argument representation and patient diagnosis, research on medical decision making used an interactive whiteboard and studied how it enabled team members to represent perspectives about data (symptoms and vital signs) as well as about solutions in the form of diagnoses (Lu et al., 2010). In problem solving research, network visualization tools have been examined as a means of promoting communications in distributed teams (Balakrishnan et al., 2008). Visualization tools support both individual and team problem solving by providing shared access to data in an externalized form (representations illustrating data as nodes). Further, these tools foster an increase in information sharing among team members that helps them better "connect the dots" and develop a shared understanding of the problem.

More recent research has explicated a catalog of action patterns and a variety of complex cognitive activities that can be utilized for technological visual representation tools to support teams (Sedig and Parsons, 2013). Studies on display design have shown how variations in externalizations influence complex collaboration. For example, a translucent interface meant to assist

sensemaking, fostered collaboration by supporting the sharing of insights and preventing narrowing of focus (Goyal and Fussell, 2016). This led to collaborators identifying more problem solving clues as well as finding a target in a criminal investigation task.

Technical domains such as architecture have also been studied to understand how externalizations foster technical work in the design process. Here, research has found that architects varied in their use of high- vs. low-resolution drawings dependent upon both task needs and the people with whom they were communicating (Retelny and Hinds, 2016). For example, architects were found to use these for both conceptual work with clients and for technical work with design teammates. In the former case, high-resolution representations supported the development of mutual understanding of the project's "design intent" as well as helped with collaborative decision making. In the case of the latter, low-resolution images would be used to provide and elicit feedback as well as resolve misinterpretations or ambiguities.

This concept has also been used in the context of scientific collaboration to show how boundary objects support interdisciplinary research. Science teams are found to create visual models and co-construct diagrams while engaged in collaborative processes (Pennington, 2010). The line of work has also integrated the idea of boundary objects with model-based reasoning to describe how scientists from different disciplines create boundary negotiating objects that support development of shared understanding (Pennington, 2011a,b). In line with early theorizing on shared problem model development (Fiore and Schooler, 2004), Pennington et al. (2016) have shown how external representations provide a firm foundation on which collaborators are able to create mutual understanding of complex problems.

In sum, boundary objects can be characterized as externalizations of cognition and may take the form of drawings, charts, graphs, prototypes, or models generated by team members, as well as tools used for project management, such as timelines and Gantt charts, or schedules and tables (see Ewenstein and Whyte, 2009). This work fits with research noting that information technology can be construed of as a form of transactive memory system (Lewis and Herndon, 2011). In the context of collaboration, the technology acts as an external memory system that is relied upon to support team processes. We suggest that boundary objects, as a technology-based form of transactive memory, can be viewed as serving the explicit purpose of facilitating coordination and collaboration between the functional or disciplinary boundaries of team members. This is particularly important for team effectiveness in that this is where common ground is not frequently held (cf. Bruns, 2013).

## Representations in Training and Learning Research

While the aforementioned research looked at externalized cognition in support of teamwork in various complex work contexts, evidence for it also comes from research on training and learning. External representations in the form of "information boards" were used in a training study of knowledge building for a collaborative planning task (Rentsch et al., 2010). Information

boards supported the creation of artifacts in the form of posts and allowed team members to organize and visually manipulate these posts. Further, it allowed them to focus shared attention on particular facets of knowledge when appropriate. Training with these artifacts supported team member transfer of knowledge and knowledge congruence, leading to overall improvements in team performance (Rentsch et al., 2010). In related research, Rentsch et al. (2014) studied training in the use of knowledge objects in collaborative problem solving. These were artifacts designed to foster schema-enriched communication in the team chats. This fostered the sharing of unique information and the transfer and congruence of knowledge across the team, leading to superior solutions.

Technology supported learning research has also been studying the externalization of cognition during collaboration. Here, visualization tools are used to externalize cognition in the form of representational artifacts such as diagrams, maps, or sketches, that help team members better understand task elements and their relations. For example, early research examined how computer support tools allow team members to jointly construct representations (Roschelle and Teasley, 1995). They found that these facilitated the definition of the problem space and the explication of executable problem solving plans. Other research has shown how computer-based visualization tools, such as matrices and graphs, are effective in helping teams learn about the connections between data, hypotheses, and evidential relationships (Suthers and Hundhausen, 2001). In a discussion of group cognition in the context of technology and learning, Stahl (2006) lays out a framework for understanding the interaction between individual and collective cognition, negotiation of meaning and understanding and how technology supports knowledge building.

Others have focused on developing technologies that help structure arguments to support learning. For example, representational tools that help information checking was found to be constructive (Kanselaar et al., 2002). Further, collaborative learning teams were shown to need help coordinating their communications as well as help in being kept on track with regard to their argumentation processes because they would often lose their thematic focus (for an early review of these tools see Kanselaar et al., 2002). Such approaches can also more specifically help teachers understand and support collaborative cognition in their classes. In an online computer science course, visualization tools were used to represent student processes and were found to help the teacher develop a better awareness of the class' performance and for students to develop self-reflection skills (Govaerts et al., 2010).

In sum, technologies that support external cognition in a learning context (e.g., 'mindtools,' visualizations, concept maps), are argued to augment knowledge acquisition by helping learners more easily represent their knowledge. With these externalizations, learners develop a shared representation that can help them transform data and information into knowledge around the content to be learned. This transformation takes place through interpretive activities, such as critical thinking or manipulative visualization, around the representations (Kirschner and Erkens, 2006).

## Summary

In sum, this review was meant to provide evidence for the external cognition perspective as it has emerged somewhat independently in a variety of domains. Although referred to with differing terms, thematically similar across these studies is the role of cognitive artifacts and various technologies supporting a variety of teamwork processes in numerous fields (see **Table 1**). Stated simply, the items reviewed above are exemplars of, and evidence for, the concept of external cognition. This is the case primarily in the sense that the required cognitive activity is distributed among members of a team and their task elements, in which artifacts serve to coordinate between internal and external structures over some duration of time (cf., Hutchins, 1995a; Hollnagel, 2002; Fiore et al., 2010b; Cooke et al., 2013; Stanton, 2016).

All this is to say that artifacts play an essential role in teamwork, and as such, leveraging the notion of external cognition to improve team process requires integrating concepts from the organizational sciences on the study of teams with relevant ideas from cognitive science. As a theoretical mechanism, then, the construct of external cognition can be conceptualized and measured as something supporting inter-related team functions (Salomon, 1993; Zhang and Norman, 1994; Zhang, 1997; Zhang and Wang, 2005; Zhang and Patel, 2006). Despite the evidence, this body of research has yet to be integrated with important concepts from organizational research on different elements of teamwork.

Toward this end, we next draw from these varied literatures, and link the findings described above to the distinction found in organizational research between "teamwork" and "taskwork." With this, we show how they can help parse different aspects of team cognition, particularly the role played by artifacts in the team environment. With a clearer description of artifacts and technological tools, coupled with concepts from team research, external cognition can be more fully integrated with team theory to understand how technology can be conceptualized and measured *as* a teammate. From this, we inform new avenues of research for team cognition to examine the ways artifacts support processes and enhance performance in the context of hybrid human technology teams.

## INTEGRATING THE ORGANIZATIONAL AND COGNITIVE SCIENCES IN THE STUDY OF TEAM COGNITION

Research in teams and team training has provided a solid foundation on which to understand and improve team process and outcomes. However, as noted, external cognition, and the role that technology plays in facilitating team processes, has yet to be fully integrated in much organizational research. In this section, we attempt to partially redress this gap by describing a framework from team theory that can be used to conceptualize and measure external team cognition, which, in turn, could inform the design of technology meant to support team performance. By adopting and adapting concepts, we contribute to team cognition research by helping to more precisely determine the role artifacts play in team process

**TABLE 1 | Representative descriptions and forms of artifacts arising in varied research literatures.**

| Area of research | General definition | Example forms | Reference |
|---|---|---|---|
| *Cognitive Artifacts in Distributed Cognition* | Cognitive artifacts are conceptualized as something constructed by humans as an aid to enhance or improve cognitive processes | Schedules, lists, display boards, patient records, digital traces, navigation technology | Hutchins, 1995a,b; Hutchins, 1999; Hollan et al., 2000; Nemeth et al., 2004, 2006; Rambusch et al., 2004; Zhang and Patel, 2006; Xiao et al., 2007 |
| *Boundary Objects in Organizational Research and Computer Supported Cooperative Work* | Boundary objects are practical artifacts that mediate interaction and shared knowledge across diverse groups and communities of practice with varying expertise and perspectives and act as a bridge for communication and coordination | Paper documents, technical drawings, whiteboards, maps, tables, computer displays, software algorithms; network diagrams, information spaces, coordinative artifacts | Star and Griesemer, 1989; Suchman and Trigg, 1991; Heath and Luff, 1992; Carlile, 2002, 2004; Yakura, 2002; Fiore and Schooler, 2004; Schmidt and Wagner, 2004; Bardram and Bossen, 2005; Levina and Vaast, 2005; Lee, 2007; Balakrishnan et al., 2008; Leonardi and Barley, 2008; Orlikowski and Scott, 2008; Ewenstein and Whyte, 2009; Leonardi, 2010, 2012; Lu et al., 2010; Star, 2010; Pennington, 2011a,b; Doolin and McLeod, 2012; Hecker, 2012; Nicolini et al., 2012; Stigliani and Ravasi, 2012; Bruns, 2013; Blomberg and Karasti, 2013; Jirotka et al., 2013; Goyal and Fussell, 2016; Retelny and Hinds, 2016 |
| *Representations and Visualization in Training and Learning Research* | Representations are forms of visualization that focus shared attention on learning elements and support development of arguments or richer knowledge structures by showing relationships across elements of the to-be-learned content | Information boards; diagrams, maps, sketches; matrices, graphs, mindtools | Roschelle and Teasley, 1995; Zhang, 1998; Suthers and Hundhausen, 2001; Kanselaar et al., 2002; Zhang and Wang, 2005; Kirschner and Erkens, 2006; Stahl, 2006; Govaerts et al., 2010; Rentsch et al., 2010, 2014 |

and how technologies mediate the creation and use of such artifacts.

## Teamwork and Taskwork in Team Cognition

The distinction between teamwork and taskwork in the organizational sciences has been a useful heuristic for conceptualizing collaboration in teams (e.g., Cannon-Bowers et al., 1995; Mathieu et al., 2000). *Teamwork* is characterized as the types of behavior essential for working together with team members including designated roles and responsibilities, interdependencies of team members, and communication patterns. *Taskwork* is characterized as the necessary functions required for meeting objectives such as the operating procedures for equipment, strategies for achieving goals, and the relationships between sub-components of a task (Mathieu et al., 2000).

Adopting this important conceptual distinction helps us to categorize the forms of external cognition detailed previously. Specifically, we can more precisely articulate clear distinctions regarding the role of externalized cognition in supporting either teamwork or taskwork. On the one hand, artifacts can support *teamwork* by providing novel and more articulated ways for understanding the workflow of the team, conveying dynamic plans, and overall, clearly displaying how the work is done (e.g., Nemeth et al., 2004; Ewenstein and Whyte, 2009) as well as facilitate communication, coordination, and shared representations across multi-disciplinary teams (Yakura, 2002). On the other hand, artifacts can support taskwork by providing novel tools for analyzing data (Suthers and Hundhausen, 2001), interpreting information (Balakrishnan et al., 2008), solving problems (Roschelle and Teasley, 1995), and making decisions (Lu et al., 2010). In short, with this distinction of teamwork and taskwork, we can illustrate *what*, specifically, the external cognition is supporting. We can take this a step further with the distinction between generic and specific competencies of teamwork and taskwork to add even greater precision for the study of team cognition.

### Generic and Specific Competencies in Teamwork and Taskwork

An additional framework from the organizational sciences that can be used to guide our understanding and measurement of external team cognition is one that explicates the team and task *competencies* necessary for successful team performance (Cannon-Bowers et al., 1995). This framework outlines how certain competencies are required in virtually all team situations, whereas others are specific to certain teams (Bowers et al., 2000). In the former, all team members need what are referred to as *team-generic competencies* regardless of the task context or the organizational setting (e.g., communication skills). In the latter, some competencies are considered to be team-specific, as they are argued to apply in only particular situations. These *team-specific competencies* are more directly related to individual teams and include knowledge of roles within the team and the abilities held by team members (Bowers et al., 2000). Relatedly, task characteristics can also be thought of along these dimensions;

namely, task-generic and task-specific competencies. Whereas *task-generic competencies* are those that are necessary across task situations (e.g., exchanging information and planning), *task-specific competencies* could include understanding the goals of a certain task or the appropriate methods for accomplishing that task.

### Integrating Teamwork/Taskwork Theory with External Cognition

This synthesis of the teamwork and taskwork concepts, combined with the notion of generic and specific team and task competencies, provides an important conceptual grounding for understanding and measuring external team cognition. By more precisely describing how artifacts and the technology used to manage them can support teams, we provide guidance on how to study team cognition within sociotechnical systems. We expect that this can be used to produce a more detailed understanding of the team processes supported by technological artifacts as they relate to the needs of specific teams as well as those that are more generic for all teams (see **Table 2**). In turn, this allows for more fine-grained theoretical specification and testing within team cognition research. Based upon this integration, and to lay the groundwork for theory development, we next provide a set of propositions for team cognition research that takes into account the role of artifacts. These are devised to unite these perspectives so as to better study teams in complex settings through a more detailed examination of the types of external cognition that a given technology can support.

- **Proposition 1 – *Context driven technology.*** We propose that the effectiveness of context driven technologies that produce externalizations are dependent upon the degree to which they are specific to both the team and the task. These should support, for example, an understanding of the team's goals and their resources. Additionally, these should help with managing whether or not the team is monitoring progress and meeting objectives.
- **Proposition 2 – *Team contingent technology.*** We propose that the effectiveness of team contingent technologies for external cognition are dependent upon the degree to which they are specific to a team, but generic to a task. These should more generally enable a team by supporting team processes like conflict resolution. However, these would do so while managing teammate specific characteristics. For example, such a system might be able to track the degree to which team members have expertise in a particular topic and leverage that knowledge to inform and guide resolutions to disagreements.

**TABLE 2 | Team and task competencies propositions for externalized cognition.**

|  |  | Relation to the task | |
|---|---|---|---|
|  |  | *Specific* | *Generic* |
| Relation to the team | *Specific* | P1. Context Driven | P2. Team Contingent |
|  | *Generic* | P3. Task Contingent | P4. Transportable |

- **Proposition 3 – *Task contingent technology*.** We propose that the effectiveness of task contingent technologies supporting externalization are dependent upon the degree to which they are specific to the task but generic as to the team. These should support the completion of particular procedures that are relevant to a large number of teams. These could support analyzing particular forms of data teams might need to make decisions. For example, in healthcare teams, these would support review of patient records, or collaborative evaluation of tests like x-rays, etc. in support of diagnosis; but any healthcare team could use them.
- **Proposition 4 – *Transportable technology*.** We propose that the effectiveness of transportable technologies for externalization are dependent upon the degree to which they are generic to a team and to a task and could support any form of team/task. These could include, for example, scheduling systems for completion of tasks or these could be a means of supporting communication processes across distributed teams. The point is that it is a more general form of technological aid that could have a broader impact on team process and performance without the need to be tailored to any particular context.

By drawing from established theory in team research, this section was meant to provide greater specificity to the team and task functions external cognition is supporting, that is, a description of *what* externalizations are supporting. In this way, we are able to better specify the role of artifacts in team process. As such, the teamwork/taskwork framework and the associated generic/specific competencies, help to conceptualize how technology can sometimes be seen as a teammate in the context of hybrid human-technology teams.

## Offloading and Scaffolding in Team Cognition

Whereas the prior section considered *what* technology and associated artifacts might support as part of team cognition, we next discuss *how* they could support cognitive processes of a team. We integrate the teamwork and taskwork dimensions with theory from the cognitive sciences to provide potential explanatory mechanisms for how artifacts developed and/or used by teams support process and performance. Specifically, the externalized view of cognition provides two constructs that help us better understand important aspects of team process and performance: *offloading* and *scaffolding* (Clark, 2008).

*Offloading* is generally the act of using the environment as a semi-permanent archive for information that can be readily available and accessed when needed, but it also used to mitigate encoding and short-term memory demands (Wilson, 2002). As such, offloading primarily serves the purpose of a memory aid that can free up cognitive resources that can then be allocated toward other team processes. In this sense, it replaces what was previously an internal form of cognitive processing such as holding an item in working memory or retrieving something from long-term memory. Often seen as an evolutionary adaptation, and by some, the center of human intelligence (Dennett, 1996), offloading fits with our points about

context and cognition reviewed earlier in that it allows for efficient utilization of the environment to reduce the complexity of memory-intensive problems (Parsell, 2006).

*Scaffolding* takes the form of externalizations of cognition that directly support team-level processes by helping to mediate and support the interaction between individual and team-level cognitive activity. Scaffolding, in this sense, supports social interaction broadly (Baron, 1991; Krueger, 2011), as well as the analysis, discussion, debate of items relevant to the team's task, and the development of the teams shared understanding (e.g., Fiore and Schooler, 2004). Specifically, technological scaffolds can help teams externalize and share knowledge by allowing for the representation and discussion of information and ideas, provide storage and access to team-level information allowing for more informed comparisons and evaluations, and act as a means for social-cognitive interaction that facilitates conversation, communication, and collaboration (McLoughlin and Luca, 2002). Further, given the virtual nature of many modern day teams, and how varied forms of technology connect such teams, scaffolding is essential for effective coordination when teams work across time and space (Fiore et al., 2003; Miles and Hollenbeck, 2013).

Indeed, it is our capability to engage in offloading and scaffolding that is what some argue to be a distinctly human trait. Further, these can be seen as a primary means through which we have made great advances in civilization because of the innovations in thinking they afford. Specifically, "our habit of offloading as much as possible of our cognitive tasks into the environment itself—extruding our mind (that is our mental projects and activities) into the surrounding world, where a host of peripheral devices we construct can store, process, and re-represent our meanings, streamlining, enhancing, and protecting the processes of transformation that are our thinking" (Dennett, 1996, pp. 134–135), has significantly expanded our cognitive capabilities beyond the limitations of our biology.

### Integrating Offloading and Scaffolding for External Team Cognition

Adding offloading and scaffolding to the team cognition literature has both theoretical and practical benefit. From the theoretical standpoint, these concepts provide a means to better understand the form of team cognition as it is emerging in complex work settings. As such, it helps us to better conceptualize *how* artifacts and technologies are enabling differing kinds of team process and/or performance outcomes (e.g., Rosen, 2010; Wiese et al., 2011). From the practical standpoint, the adoption and adaptation of these constructs from the cognitive sciences will provide greater precision in, and mechanisms for, measuring external team cognition.

To guide examination of the relation between team cognition and technological artifacts, we provide the following research questions for assessing external team cognition as a means of offloading or as scaffolding. These are provided to show how theorizing from cognitive science can help lay the groundwork for research on technological supports designed to improve process and performance of teams as sociotechnical systems (see **Table 3**).

**TABLE 3 | Framework for guiding research on external cognition.**

| Focus of support | Role of external cognition | |
| --- | --- | --- |
| | *Offloading* | *Scaffolding* |
| *Taskwork* | R.Q. 1.1 and 1.2 | R.Q. 3.1 and 3.2 |
| *Teamwork* | R.Q. 2.1 and 2.2 | R.Q. 4.1 and 4.2 |

1. **Are technologies providing externalizations supporting *taskwork* through offloading?**

   - *Research Question 1.1.* Are technologies devised to support memory based elements of the task (e.g., storage of operating procedures) being relied upon by team members to meet their objectives (e.g., accessing a manual for trouble shooting)?
   - *Research Question 1.2.* Are technologies devised as repositories for task relevant data (e.g., bulleted lists of relevant information for later use) being used by team members to meet task needs?

2. **Are technologies providing externalizations supporting *teamwork* through offloading?**

   - *Research Question 2.1.* Are technologies devised to support understanding of team process (e.g., graphical representation of workflows across members) being used by the team to meet their objectives?
   - *Research Question 2.2.* Are technologies devised in support of communication (e.g., mapping who knows what and related role interdependencies) being relied upon as a guide for team related knowledge?

3. **Are technologies providing externalizations supporting *taskwork* through scaffolding?**

   - *Research Question 3.1.* Are technologies devised to support task-relevant activities (e.g., dynamically updated visualizations to guide information interpretation) being used by team members as they interact to work toward goals?
   - *Research Question 3.2.* Are technologies devised to support understanding of task-related elements (e.g., visualizations helping to illustrate relations between data), being utilized as they collaborate to meet objectives?

4. **Are technologies providing externalizations supporting *teamwork* through scaffolding?**

   - *Research Question 4.1.* Are technologies devised to support interaction processes (e.g., helping to represent ideas around arguments to foster constructive conflict) being relied upon in service of their teamwork?
   - *Research Question 4.2.* Are technologies devised to support team outcomes (e.g., helping teams evaluate solution alternatives to reach consensus) being used to meet performance outcomes?

In sum, we have provided this representative set of questions in such a way that researchers can see how to integrate the concepts of offloading and scaffolding with their own theorizing on team cognition. By framing these within the context of teamwork and taskwork as well as offloading and scaffolding, we offer theoretical concepts that, themselves, can augment existing theory. In this way, sociotechnical systems research can make significant strides in understanding and explaining the ways in which artifacts, and the technologies supporting their development and use, can be construed of as part of a larger system that is, essentially, a team-technology hybrid.

## DISCUSSION

Our goal with this paper was twofold. First, we set out to provide an overview of the externalized view of cognition in the context of team process and performance. We add to theory that views cognition within individuals and within teams as something spanning team members and their technology (Hutchins, 1995a; Hollnagel, 2002; Stanton et al., 2006; Fiore et al., 2010b; Cooke et al., 2013). Second, we lay the groundwork for future research to consider and measure this type of cognition as it occurs across individuals, team members, technologies, and artifacts. We expect that adopting this approach will lead to an enriched perspective of team cognition theory that will augment many lines of research and measurement methods. We have provided preliminary progress toward this by integrating ideas on external cognition with insights from varying disciplines that have, so far, shown little integration. Specifically, on the one hand, we have drawn from team theory in the organizational sciences to articulate how research can examine *what* and *where* technology and artifacts support team process and performance; that is, teamwork and taskwork. On the other hand, we have drawn from the cognitive sciences to articulate the ways research can examine *how* technology and artifacts support team process and performance; that is, offloading and scaffolding.

Through this integration, we are able to connect related concepts from across a disparate set of disciplines. Foundational to this was the need to illustrate how team cognition researchers could leverage ideas emerging from fields ranging from cognitive engineering, to computer supported collaborative work, to the organizational and cognitive sciences. Specifically, by blending theory from research on teamwork (Cannon-Bowers et al., 1995), with concepts from the organizational sciences (e.g., Carlile, 2004; Hecker, 2012), the cognitive sciences (e.g., Zhang and Norman, 1994; Clark, 2001a), and cognitive engineering (Hollnagel, 2002; Fiore et al., 2010b), we provide a framework for understanding and measuring how team process and performance is altered through the use of artifacts and technology. This was framed within recent theorizing on MiTs, which takes into account the role of artifacts in both internalized and externalized cognition (Fiore et al., 2008, 2010b,c), as well as theorizing on extended cognition (Clark and Chalmers, 1998; Clark, 2001a).

Note that this approach is different from others who have construed of teams *as* technology (Wallace and Hinsz, 2010). In that line of theorizing, teams are, themselves, viewed as a form of technology that is used to transform internalized cognitive resources into team solutions. Likewise, while we share similar views on team cognition with ITC theory (Cooke et al., 2013; Cooke, 2015), and DSA (e.g., Stanton et al., 2006), our focus and contribution here is distinct. That is, we elaborate upon, and extend, such efforts by making explicit the systemic relationship between team cognitive processes and the types of technological artifacts that facilitate both the externalization of knowledge and effective team performance. Our approach adds to such thinking in that we broaden what role technology potentially plays in team cognition and, indeed, should be seen as a fundamental element of the team.

As such, our efforts support recent calls by those in the organizational sciences to develop a richer understanding of the modern workplace and the complexities inherent given the role of technology in business processes (Juillerat, 2010; Bell and Kozlowski, 2012; Stigliani and Ravasi, 2012; Kozlowski et al., 2015). Further, this framework supports recent work in the study of scientific collaboration and the interaction of people and technology in support of innovation (e.g., Fiore, 2008; Asencio et al., 2012; Cummings et al., 2013).

An additional implication of the integration we have provided is that it can help to broaden current understandings of team cognition and how it is conceptualized and measured. Meta-analytic studies have shown how aspects of team cognition (e.g., shared mental models) are predictive of team process (DeChurch and Mesmer-Magnus, 2010b; Turner et al., 2014) and furthered our understanding of how compositional and compilational variables relate to team process and performance (DeChurch and Mesmer-Magnus, 2010a). We have provided a more precise framework for understanding the form and role of technological artifacts in team cognition. This broadens our conceptualization of what can be part of a shared mental model and/or what role technology plays in transactive memory systems (Austin, 2003; Lewis, 2003; Zhang et al., 2007; Huber and Lewis, 2010; Lewis and Herndon, 2011; Tollefsen et al., 2013), and cross-disciplinary coordination and collaboration (e.g., Susi et al., 2003; Gittell and Weiss, 2004; Gittell, 2006; Rico et al., 2008; Okhuysen and Bechky, 2009). These ideas can also fit within new methods for measuring teams, such as social network analysis (e.g., Leenders et al., 2016). For example, it is possible to see how artifacts utilized by teams can be viewed as nodes in a network that are part of collaboration. Thus, our accounting provides guidance on future efforts to advance the science of teams in complex sociotechnical settings by detailing additional factors for studying mediation and moderation in meta-analytic work on team cognition and suggests ways to inform the design of new tools for enhancing both team process and performance.

Viewed more broadly, our focus can be seen as an argument that technology, broadly construed, needs to be taken more seriously as a member of a team. The conceptual frameworks we put forth in this regard, are timely in that technology is going to play an increasing role in teams. We are now seeing semi-autonomous robots as members of teams in complex and high-stakes environments. For example, the military is making use of them in settings such as explosive detonation whereas, in civilian settings, robots are playing a significant role in areas such as search and rescue. Furthermore, organizations will soon be confronted with the reality of intelligent technology, in various forms, in the workplace. Whether this be in the form of cognitive computing and artificially intelligent support systems (e.g., stock trading; medical diagnosis), or embodied robots interacting on the factory floor, autonomous systems will need to be studied by organizational scientists. We provide a foundation on which to more fully examine how these systems will function as a member of a team and provide conceptual grounding for the next evolution of team research.

In this context, our integration provides a foundation for the next phase of team cognition research, a phase that will increasingly be studying hybrid human-technology teams (e.g., Wiltshire and Fiore, 2014). We provide a scaffold for understanding, not just how humans draw from, and rely on, technology in the context of teams. We additionally provide a foundation for the coming infusion of new technologies (e.g., cognitive computing, robotics) in organizational settings. The prevalence of artificial intelligence in computing systems (e.g., decision support), and machine production (e.g., industrial robotics), will only become more commonplace in the workplace. Because of this, researchers in team cognition need to have the conceptual scaffolds that well help link these technology developments to their theorizing and increase our understanding of team effectiveness.

Further, the current framework calls for, and contributes to, new forms of interdisciplinary research in the study of teams by helping to develop additional ways to conceptualize and measure team cognition. As the sophistication of technology continues to advance, and as humans continue to integrate these advances in their lives, we must, ourselves, become more sophisticated in how we study these phenomena. As Clark (2001a) articulated so well, collaboration between humans and technology should be viewed as a *continuous reciprocal causation*; specifically:

> Much of what matters about human intelligence is hidden not in the brain, nor in the technology, but in the complex and iterated interactions and collaborations between the two.... The study of these interaction spaces is not easy, and depends both on new multidisciplinary alliances and new forms of modeling and analysis. The pay-off, however, could be spectacular: nothing less than a new kind of cognitive collaborative collaboration involving neuroscience, physiology, and social, cultural, and technological studies (Clark, 2001a, p. 154).

We further this view to suggest the need for an externalized view of cognition as it relates to teams and their associated teamwork and taskwork. In doing so, external cognition provides a means for enriching study of the interdependencies across both individuals and teams and their use of artifacts and technologies such that the team competencies required for

effective performance can be more fully examined. Further, when these complementary distinctions are integrated, this can better inform our understanding of the role of technological support systems in team cognition.

## CONCLUSION

For researchers in sociotechnical systems, we emphasize that an interdisciplinary collaboration between the cognitive, organizational, and computational sciences is needed. Such research would, not only be aimed at understanding and enhancing team process and performance, but would also serve the design and delivery of approaches that better support teams in many of society's current, and future, complex socio-technical systems.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## REFERENCES

Anderson, M. L., Richardson, M. J., and Chemero, A. (2012). Eroding the boundaries of cognition: implications of embodiment. *Top. Cogn. Sci.* 4, 717–730. doi: 10.1111/j.1756-8765.2012.01211.x

Asencio, R., Carter, D. R., DeChurch, L. A., Zaccaro, S. J., and Fiore, S. M. (2012). Charting a course for collaboration: a multiteam perspective. *Transl. Behav. Med. Pract. Policy Res.* 4, 487–494. doi: 10.1007/s13142-012-0170-3

Austin, J. R. (2003). Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. *J. Appl. Psychol.* 88, 866–878. doi: 10.1037/0021-9010.88.5.866

Balakrishnan, A. D., Fussell, S. R., and Kiesler, S. (2008). "Do visualization improve synchronous remote collaboration?," in *Proceedings of ACM CHI: Conference on Human Factors in Computing Systems*, San Jose, CA, 1227–1236.

Bardram, J. E., and Bossen, C. (2005). "A web of coordinative artifacts: collaborative work at a hospital ward," in *Proceedings of the International Conference on Supporting Group Work (GROUP'05)*, (New York: ACM Press), 168–176.

Baron, L. (1991). Peer tutoring, microcomputer learning and young children. *J. Comput. Child. Educ.* 2, 27–40.

Bell, B. S., and Kozlowski, S. W. (2012). Three conceptual themes for future research on teams. *Ind. Organ. Psychol.* 5, 45–48. doi: 10.1111/j.1754-9434.2011.01403.x

Blomberg, J., and Karasti, H. (2013). Reflections on 25 years of ethnography in CSCW. *Comput. Support. Coop. Work* 22, 373–423. doi: 10.1007/s10606-012-9183-1

Bourbousson, J., Poizat, G., Saury, J., and Seve, C. (2011). Description of dynamic shared knowledge: an exploratory study during a competitive team sports interaction. *Ergonomics* 54, 120–138. doi: 10.1080/00140139.2010.544763

Bowers, C. A., Jentsch, F., and Salas, E. (2000). "Establishing aircrew competencies: a comprehensive approach for identifying CRM training needs," in *Aircrew Training and Assessment*, eds H. F. O'Neil and D. Andrews (Mahwah, NJ: Erlbaum), 67–84.

Bowers, C. E., Salas, E. E., and Jentsch, F. E. (eds). (2006). *Creating High-Tech Teams: Practical Guidance on Work Performance and Technology*. Washington, DC: American Psychological Association.

Brodbeck, F. C., Kerschreiter, R., Mojzisch, A., and Schulz-Hardt, S. (2007). Group decision making under conditions of distributed knowledge: the information asymmetries model. *Acad. Manag. Rev.* 32, 459–479. doi: 10.5465/AMR.2007.24351441

Bruns, H. C. (2013). Working alone together: coordination in collaboration across domains of expertise. *Acad. Manag. J.* 56, 62–83. doi: 10.5465/amj.2010.0756

Cacciabue, P. C., and Hollnagel, E. (1995). "Simulation of cognition: applications," in *Expertise and Technology: Cognition and Human-Computer Cooperation*, eds J. M. Hoc, P. C. Cacciabue, and E. Hollnagel (Denmark: Lawrence Erlbaum Associates), 55–73.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., and Volpe, C. E. (1995). "Defining competencies and establishing team training requirements," in *Team Effectiveness and Decision Making in Organizations*, eds R. A. Guzzo and E. Salas (San Francisco, CA: Jossey-Bass), 333–381.

Carlile, P. R. (2002). A pragmatic view of knowledge and boundaries: boundary objects in new product development. *Organ. Sci.* 13, 442–455. doi: 10.1287/orsc.13.4.442.2953

Carlile, P. R. (2004). Transferring, translating, and transforming: an integrative framework for managing knowledge across boundaries. *Organ. Sci.* 15, 555–568. doi: 10.1287/orsc.1040.0094

Clancey, W. J. (1997). *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge: Cambridge University Press.

Clark, A. (2001a). *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford: Oxford University Press.

Clark, A. (2001b). Reasons, robots, and the extended mind. *Mind Lang.* 15, 121–145. doi: 10.1111/1468-0017.00162

Clark, A. (2002). Towards a science of the bio-technological mind. *Int. J. Cogn. Technol.* 1, 21–33. doi: 10.1075/ijct.1.1.03cla

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7

Cooke, N. J. (2015). Team cognition as interaction. *Curr. Dir. Psychol. Sci.* 24, 415–419. doi: 10.1177/0963721415602474

Cooke, N. J., and Fiore, S. M. (2009). "Cognitive science based principles for the design and delivery of training," in *Learning, Training, and Development in Organizations*, eds S. W. J. Kozlowski and E. Salas (Mahwah, NJ: Erlbaum), 169–201.

Cooke, N. J., and Gorman, J. C. (2009). Interaction-based measures of cognitive systems. *J. Cogn. Eng. Decis. Mak.* 3, 27–46. doi: 10.1518/155534309X433302

Cooke, N. J., Gorman, J. C., and Kiekel, P. A. (2008). "Communication as team-level cognitive processing," in *Macrocognition in Teams: Theories and Methodologies*, eds M. Letsky, N. Warner, S. M. Fiore, and C. A. P. Smith (Hants: Ashgate Publishing Ltd), 51–64.

Cooke, N. J., Gorman, J. C., Myers, C. W., and Duran, J. L. (2013). Interactive team cognition. *Cogn. Sci.* 37, 255–285. doi: 10.1111/cogs.12009

Cooke, N. J., Gorman, J. C., and Rowe, L. J. (2009). "An ecological perspective on team cognition," in *Team Effectiveness in Complex Organizations: Cross-Disciplinary Perspectives and Approaches*, eds E. Salas, J. Goodwin, and C. S. Burke (New York, NY: Taylor & Francis), 157–182.

Cooke, N. J., Kiekel, P. A., and Helm, E. E. (2001). Measuring team knowledge during skill acquisition of a complex task. *Int. J. Cogn. Ergon.* 5, 297–315. doi: 10.1207/S15327566IJCE0503_10

Cooke, N. J., Salas, E., Kiekel, P. A., and Bell, B. (2004). "Advances in measuring team cognition," in *Team Cognition: Understanding the Factors that Drive Process and Performance*, eds E. Salas and S. M. Fiore (Washington, DC: American Psychological Association), 83–106.

Cummings, J. N., Kiesler, S., Bosagh Zadeh, R., and Balakrishnan, A. D. (2013). Group heterogeneity increases the risks of large group size: a longitudinal study of productivity in research groups. *Psychol. Sci.* 24, 880–890. doi: 10.1177/0956797612463082

de Vries, E., and Masclet, C. (2013). A framework for the study of external representations in collaborative design settings. *Int. J. Hum. Comput. Stud.* 71, 46–58. doi: 10.1016/j.ijhcs.2012.07.005

DeChurch, L. A., and Mesmer-Magnus, J. R. (2010a). Measuring shared team mental models: a meta-analysis. *Group Dyn.* 14, 1–14. doi: 10.1037/a0017455

DeChurch, L. A., and Mesmer-Magnus, J. R. (2010b). The cognitive underpinnings of team effectiveness: a meta-analysis. *J. Appl. Psychol.* 95, 32–53. doi: 10.1037/a0017328

Dennett, D. C. (1996). *Kinds of Minds*. New York, NY: Basic Books.

Doolin, B., and McLeod, L. (2012). Sociomateriality and boundary objects in information systems development. *Eur. J. Inform. Syst.* 21, 570–586. doi: 10.1057/ejis.2012.20

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Hum. Factors* 37, 32–64. doi: 10.1518/001872095779049543

Espinosa, J. A., Lerch, F. J., and Kraut, R. E. (2004). "Explicit versus implicit coordination mechanisms and task dependencies: one size does not fit all," in *Team Cognition: Understanding the Factors that Drive Process and Performance*, eds E. Salas and S. M. Fiore (Washington, DC: American Psychological Association), 107–129.

Espinosa, J. A., Slaughter, S. A., Kraut, R. E., and Herbsleb, J. D. (2007). Team knowledge and coordination in geographically distributed software development. *J. Manag. Inform. Syst.* 24, 135–169. doi: 10.2753/MIS0742-1222240104

Ewenstein, B., and Whyte, J. (2009). Knowledge practices in design: the role of visual representations as 'Epistemic Objects'. *Organ. Stud.* 30, 7–30. doi: 10.1177/0170840608083014

Faraj, S., and Sproull, L. (2000). Coordinating expertise in software development teams. *Manag. Sci.* 46, 1554–1568. doi: 10.1287/mnsc.46.12.1554.12072

Fioratou, E., Chatzimichailidou, M. M., Grant, S., Glavin, R., Flin, R., and Trotter, C. (2016). Beyond monitors: distributed situation awareness in anaesthesia management. *Theor. Issues Ergon. Sci.* 17, 104–124. doi: 10.1080/1463922X.2015.1106620

Fiore, S. (2013). "Situated action," in *Encyclopedia of Philosophy and the Social Sciences*, Vol. 18, ed. B. Kaldis (Thousand Oaks, CA: SAGE Publications), 869–871.

Fiore, S. M. (2008). Interdisciplinarity as teamwork: how the science of teams can inform team science. *Small Group Res.* 39, 251–277. doi: 10.1177/1046496408317797

Fiore, S. M. (2012). "Cognition and technology: interdisciplinarity and the impact of cognitive engineering research on organizational productivity," in *Oxford Handbook of Industrial and Organizational Psychology*, ed. S. Koslowski (Oxford: Oxford University Press), 1306–1322.

Fiore, S. M., Elias, J., Salas, E., Warner, N., and Letsky, M. (2010a). "From data, to information, to knowledge: measuring knowledge building in the context of collaborative cognition," in *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams*, eds E. Patterson and C. Miller (Farnham: Ashgate Publishing), 179–200.

Fiore, S. M., Rosen, M., Salas, E., Burke, S., and Jentsch, F. (2008). "Processes in complex team problem solving: parsing and defining the theoretical problem space," in *Macrocognition in Teams: Theories and Methodologies*, eds M. Letsky, N. Warner, S. M. Fiore, and C. Smith (London: Ashgate Publishers), 143–163.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., and Warner, N. (2010b). Toward an understanding of macrocognition in teams: predicting processes in complex collaborative contexts. *Hum. Factors* 52, 203–224. doi: 10.1177/0018720810369807

Fiore, S. M., Salas, E., Cuevas, H. M., and Bowers, C. (2003). Distributed coordination space: towards a theory of distributed team process and performance. *Theor. Issues Ergon. Sci.* 4, 340–364. doi: 10.1080/1463922021000049971

Fiore, S. M., and Schooler, J. W. (2004). "Process mapping and shared cognition: teamwork and the development of shared problem models," in *Team Cognition: Understanding the Factors that Drive Process and Performance*, eds E. Salas and S. M. Fiore (Washington, DC: American Psychological Association), 133–152.

Fiore, S. M., Smith-Jentsch, K. A., Salas, E., Warner, N., and Letsky, M. (2010c). Toward an understanding of macrocognition in teams: developing and defining complex collaborative processes and products. *Theor. Issues Ergon. Sci.* 11, 250–271. doi: 10.1080/14639221003729128

Gittell, J. H. (2006). "Relational coordination: coordinating work through relationships of shared goals, shared knowledge and mutual respect," in *Relational Perspectives in Organizational Studies: A Research Companion*, eds O. Kyriakidou and M. Ozbilgin (Cheltenham, PA: Edward Elgar Publishers).

Gittell, J. H., and Weiss, L. (2004). Coordination networks within and across organizations: a multi-level framework. *J. Manag. Stud.* 41, 127–153. doi: 10.1186/1472-6963-13-296

Gorman, J. C. (2014). Team coordination and dynamics: two central issues. *Curr. Dir. Psychol. Sci.* 23, 355–360. doi: 10.1177/0963721414545215

Gorman, J. C., Cooke, N. J., and Amazeen, P. G. (2010). Training adaptive teams. *Hum. Factors* 52, 295–307. doi: 10.1177/0018720810371689

Govaerts, S., Verbert, K., Klerkx, J., and Duval, E. (2010). "Visualizing activities for self-reflection and awareness," in *Advances in Web-Based Learning–ICWL 2010*, Vol. 6483, *Lecture of Notes in Computer Science*, eds X. Luo, M. Spaniol, L. Wang, Q. Li, W. Nejdl, and W. Zhang (Berlin: Springer), 91–100.

Goyal, N., and Fussell, S. R. (2016). "Effects of Sensemaking Translucence on Distributed Collaborative Analysis," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing*, (New York, NY: ACM), 288–302.

Guzzo, R. A., and Dickson, M. W. (1996). Teams in organizations: recent research on performance and effectiveness. *Annu. Rev. Psychol.* 47, 307–338. doi: 10.1146/annurev.psych.47.1.307

Heath, C., and Luff, P. (1992). Collaboration and control: crisis management and multimedia technology in london underground line control rooms. *Comput. Support. Coop. Work* 1, 69–94. doi: 10.1007/BF00752451

Hecker, A. (2012). Knowledge beyond the individual? Making sense of a notion of collective knowledge in organization theory. *Organ. Stud.* 33, 423–445. doi: 10.1177/0170840611433995

Hinsz, V. B., Tindale, R. S., and Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychol. Bull.* 121, 43–64. doi: 10.1037/0033-2909.121.1.43

Hinsz, V. B., Vollrath, D. A., Nagao, D. H., and Davis, J. H. (1988). Comparing the structure of individual and small group perceptions. *Small Group Res.* 4, 159–168.

Hoeft, R. M., Kochan, J. A., and Jentsch, F. (2006). "Automated systems in the cockpit: is the autopilot, "George," a team member?," in *Creating High-Tech Teams: Practical Guidance on Work Performance and Technology*, eds C. Bowers, E. Salas, and F. Jentsch (Washington, DC: American Psychological Association), 243–259.

Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput. Hum. Interact.* 7, 174–196. doi: 10.1145/353485.353487

Hollnagel, E. (2002). Cognition as control: a pragmatic approach to the modeling of joint cognitive systems. *Theor. Issues Ergon. Sci.* 2, 309–315. doi: 10.1080/14639220110104934

Huber, G. P., and Lewis, K. (2010). Cross-understanding: implications for group cognition and performance. *Acad. Manag. Rev.* 35, 6–26. doi: 10.5465/AMR.2010.45577787

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

Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cogn. Sci.* 19, 265–288. doi: 10.1207/s15516709cog1903_1

Hutchins, E. (1999). "Cognitive artifacts," in *MIT Encyclopedia of the Cognitive Sciences*, ed. F. C. Keil Cambridge, MA: MIT Press.

Jenkins, M. P., Gross, G., Bisantz, A. M., and Nagi, R. (2011). "Towards context-aware hard/soft information fusion: incorporating situationally qualified human observations into a fusion process for intelligence analysis," in *Proceedings of the 2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)* (Rome: IEEE), 74–81.

Jirotka, M., Lee, C. P., and Olson, G. M. (2013). Supporting scientific collaboration: methods, tools and concepts. *Comput. Support. Coop. Work* 22, 667–715. doi: 10.1016/j.gerinurse.2010.05.004

Juillerat, T. L. (2010). Friends, not foes?: work design and formalization in the modern work context. *J. Organ. Behav.* 31, 216–239. doi: 10.1002/job.654

Kanselaar, G., Erkens, G., Andriessen, J., Prangsma, M., Veerman, A., and Jaspers, J. (2002). "Designing argumentation tools for collaborative learning," in *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*, eds P. A. Kirschner, S. J. Shum, and C. S. Carr (London: Springer-Verlag), 51–74.

Kiekel, P. A., and Cooke, N. J. (2004). "Human factors aspects of team cognition," in *The Handbook of Human Factors in Web Design*, (Mahwah, NJ: Lawrence Erlbaum), 90–103.

Kirschner, P. A., and Erkens, G. (2006). Cognitive tools and mindtools for collaborative learning. *J. Educ. Comput. Res.* 35, 199–209. doi: 10.2190/R783-230M-0052-G843

Kitchin, J., and Baber, C. (2016). A comparison of shared and distributed situation awareness in teams through the use of agent-based modelling. *Theor. Issues Ergon. Sci.* 17, 8–41. doi: 10.1080/1463922X.2015.1106616

Klein, G., Moon, B., and Hoffman, R. R. (2006). Making sense of sensemaking 2: a macrocognitive model. *IEEE Intell. Syst.* 21, 88–92. doi: 10.1109/MIS.2006.75

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., and Hollnagel, E. (2003). Macrocognition. *IEEE Intell. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Kozlowski, S. W., Grand, J. A., Baard, S. K., and Pearce, M. (2015). "Teams, teamwork, and team effectiveness: implications for human systems integration," in *APA Handbook of Human Systems Integration*, eds D. Boehm-Davis, F. T. Durso, and J. D. Lee (Washington, DC: American Psychological Association), 555–571.

Krueger, J. (2011). Extended cognition and the space of social interaction. *Conscious. Cogn.* 20, 643–657. doi: 10.1016/j.concog.2010.09.022

Larkin, J. H., and Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* 11, 65–100. doi: 10.1111/j.1551-6708.1987.tb00863.x

Larson, J. R., and Christensen, C. (1993). Groups as problem-solving units: toward a new meaning of social cognition. *Br. J. Soc. Psychol.* 32, 5–30. doi: 10.1111/j.2044-8309.1993.tb00983.x

Lee, C. P. (2007). Boundary negotiating artifacts: unbinding the routine of boundary objects and embracing chaos in collaborative work. *Comput. Support. Coop. Work* 16, 307–339. doi: 10.1007/s10606-007-9044-5

Lee, J. D., and Kirlik, A. (eds). (2013). *The Oxford Handbook of Cognitive Engineering*. Oxford: Oxford University Press.

Leenders, R. T. A., Contractor, N. S., and DeChurch, L. A. (2016). Once upon a time understanding team processes as relational event networks. *Organ. Psychol. Rev.* 6, 92–115. doi: 10.1177/2041386615578312

Leonardi, P. M. (2010). Digital materiality? How artifacts without matter, matter. *First Monday* 15. doi: 10.5210/fm.v15i6.3036

Leonardi, P. M. (2012). "Materiality, sociomateriality, and socio-technical systems: what do these terms mean? how are they different? do we need them?," in *Materiality and Organizing: Social Interaction in a Technological World*, eds P. M. Leonardi, B. A. Nardi, and J. Kallinikos (Oxford: Oxford University Press), 25–48.

Leonardi, P. M., and Barley, S. R. (2008). Materiality and change: challenges to building better theory about technology and organizing. *Inform. Organ.* 18, 159–176. doi: 10.1016/j.infoandorg.2008.03.001

Letsky, M., Warner, N., Fiore, S. M., and Smith, C. (eds). (2008). *Macrocognition in Teams: Theories and Methodologies*. London: Ashgate.

Levina, N., and Vaast, E. (2005). The emergence of boundary spanning competence in practice: implications for implementation and use of information systems. *MIS Q.* 29, 335–363.

Lewis, K. (2003). Measuring transactive memory systems in the field: scale development and validation. *J. Appl. Psychol.* 88, 587–604. doi: 10.1037/0021-9010.88.4.587

Lewis, K., and Herndon, B. (2011). Transactive memory systems: current issues and future research directions. *Organ. Sci.* 22, 1254–1265. doi: 10.1287/orsc.1110.0647

Lord, R. G., and Maher, K. M. (1991). "Cognitive theory in industrial and organizational psychology," in *Handbook of Industrial Organizational Psychology*, 2nd Edn, Vol. 2, eds M. Dunnette and L. Hough (Palo Alto, CA: Consulting Psychological Press), 1–62.

Lu, J., Lajoie, S. P., and Wiseman, J. (2010). Scaffolding problem-based learning with CSCL tools. *Comput. Support. Collab. Learn.* 5, 283–298. doi: 10.1007/s11412-010-9092-6

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *J. Appl. Psychol.* 35, 273–283. doi: 10.1037/0021-9010.85.2.273

McLoughlin, C., and Luca, J. (2002). A learner–centered approach to developing team skills through web–based learning and assessment. *Br. J. Educ. Technol.* 33, 571–582. doi: 10.1111/1467-8535.00292

Miles, J., and Hollenbeck, J. R. (2013). "Teams and technology," in *Psychology of Workplace Technology*, eds M. D. Coovert and L. F. Thompson (London: Taylor & Francis), 99–117.

Mitchell, R. K., Randolph-Seng, B., and Mitchell, J. R. (2011). Socially situated cognition: imagining new opportunities for entrepreneurship research. *Acad. Manag. Rev.* 36, 774–776. doi: 10.5465/amr.2011.0001

Nemeth, C., O'Connor, M., Klock, P. A., and Cook, R. (2006). Discovering healthcare cognition: the use of cognitive artifacts to reveal cognitive work. *Organ. Stud.* 2, 1011–1035. doi: 10.1177/0170840606065708

Nemeth, C. P., Cook, R. I., O'Connor, M., and Klock, P. A. (2004). Using cognitive artifacts to understand distributed cognition. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* 34, 726–735. doi: 10.1109/TSMCA.2004.836798

Nicolini, D., Mengis, J., and Swan, J. (2012). Understanding the role of objects in cross-disciplinary collaboration. *Organ. Sci.* 23, 612–629. doi: 10.1287/orsc.1110.0664

Norman, D. A. (1991). "Cognitive artifacts," in *Designing Interaction: Psychology at the Human-Computer Interface*, ed. J. M. Carroll (Cambridge: Cambridge University Press), 17–38.

Okhuysen, G. A., and Bechky, B. A. (2009). Coordination in organizations: an integrative perspective. *Acad. Manag. Ann.* 3, 463–502. doi: 10.1080/19416520903047533

Orlikowski, W. J., and Scott, S. V. (2008). Sociomateriality: challenging the separation of technology, work and organization. *Acad. Manag. Ann.* 2, 433–474. doi: 10.1080/19416520802211644

Parsell, M. (2006). The cognitive cost of extending an evolutionary mind into the environment. *Cogn. Process.* 7, 3–10. doi: 10.1007/s10339-005-0015-9

Pennathur, P. R., Guarrera, T. K., Bisantz, A. M., Fairbanks, R. J., Perry, S. J., and Wears, R. L. (2008). "Cognitive artifacts in transition: an analysis of information content changes between manual and electronic patient tracking systems," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 52, (Thousand Oaks, CA: SAGE Publications), 363–367.

Pennington, D. (2010). The dynamics of material artifacts in collaborative research teams. *Comput. Support. Coop. Work* 19, 175–199. doi: 10.1007/s10606-010-9108-9

Pennington, D. D. (2011a). Bridging the disciplinary divide: co-creating research ideas in eScience teams. *Comput. Support. Coop. Work* 20, 165–196. doi: 10.1007/s10606-011-9134-2

Pennington, D. D. (2011b). Collaborative, cross-disciplinary learning and co-emergent innovation in eScience teams. *Earth Sci. Inform.* 4, 55–68. doi: 10.1007/s12145-011-0077-4

Pennington, D. D., Bammer, G., Danielson, A., Gosselin, D., Gouvea, J., Habron, G., et al. (2016). The EMBeRS project: employing model-based reasoning in socio-environmental synthesis. *J. Environ. Stud. Sci.* 6, 278–286. doi: 10.1007/s13412-015-0335-8

Rambusch, J., Susi, T., and Ziemke, T. (2004). "Artefacts as mediators of distributed social cognition: a case study," in *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, (Mahwah, NJ: Lawrence Erlbaum), 1113–1118.

Rentsch, J. R., Delise, L. A., Mello, A. L., and Staniewicz, M. J. (2014). The integrative team knowledge building strategy in distributed problem-solving teams. *Small Group Res.* 45, 568–591. doi: 10.1177/1046496414537690

Rentsch, J. R., Delise, L. A., Salas, E., and Letsky, M. P. (2010). Facilitating knowledge building in teams: can a new team training strategy help? *Small Group Res.* 41, 505–523. doi: 10.1177/1046496410369563

Retelny, D., and Hinds, P. (2016). "Embedding intentions in drawings: how architects craft and curate drawings to achieve their goals," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, (New York, NY: ACM), 1310–1322.

Rico, R., Sánchez-Manzanares, M., Gil, F., and Gibson, C. (2008). Team implicit coordination processes: a team knowledge–based approach. *Acad. Manag. Rev.* 33, 163–184. doi: 10.5465/AMR.2008.27751276

Rico, R., Sánchez-Manzanares, M., and Gil, F. (2011). Coordination processes in work teams. *Papeles del Psicólogo* 32, 59–68.

Roschelle, J., and Teasley, S. D. (1995). "The construction of shared knowledge in collaborative problem solving," in *Computer Supported Collaborative Learning*, ed. C. O'Malley (Berlin: Springer-Berlag), 69–97.

Rosen, M. A. (2010). *Collaborative Problem Solving: The Role of Team Knowledge Building Processes and External Representations*. Ph.D. Dissertation, University of Central Florida, Orlando, FL.

Salas, E., and Fiore, S. M. (eds). (2004). *Team Cognition: Understanding the Factors that Drive Process and Performance*. Washington, DC: American Psychological Association.

Salas, E., Fiore, S. M., and Letsky, M. P. (eds). (2012). *Theories of Team Cognition: Cross-Disciplinary Perspectives*. New York, NY: Taylor & Francis Group.

Salomon, G. (1993). "No distribution without individuals' cognition: a dynamic interactional view," in *Distributed Cognitions: Psychological and Educational Considerations*, ed. G. Salomon (New York, NY: Cambridge University Press), 111–138.

Schmidt, K., and Wagner, I. (2004). Ordering systems: coordinative practices and artifacts in architectural design and planning. *Comput. Support. Coop. Work* 13, 349–408. doi: 10.1007/s10606-004-5059-3

Sedig, K., and Parsons, P. (2013). Interaction design for complex cognitive activities with visual representations: a pattern-based approach. *AIS Trans. Hum. Comput. Interact.* 5, 84–133.

Sorensen, L. J., and Stanton, N. A. (2013). Y is best: how distributed situational awareness is mediated by organisational structure and correlated with task success. *Saf. Sci.* 56, 72–79. doi: 10.1016/j.ssci.2012.05.026

Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: MIT Press.

Stanton, N. A. (2016). Distributed situation awareness. *Theor. Issues Ergon. Sci.* 17, 1–7. doi: 10.1080/1463922X.2015.1106615

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., et al. (2006). Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics* 49, 1288–1311. doi: 10.1080/00140130600612762

Star, S. L. (2010). This is not a boundary object: reflections on the origin of a concept. *Sci. Technol. Hum. Values* 35, 601–617. doi: 10.1179/2046905514Y.0000000146

Star, S. L., and Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Soc. Stud. Sci.* 19, 387–420. doi: 10.1177/030631289019003001

Stigliani, I., and Ravasi, D. (2012). Organizing thoughts and connecting brains: material Practices and the transition from individual to group-level prospective sensemaking. *Acad. Manag. J.* 55, 1232–1259. doi: 10.5465/amj.2010.0890

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human-machine Communication*. Cambridge: Cambridge University Press.

Suchman, L. (2007). *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press.

Suchman, L. A., and Trigg, R. H. (1991). "Understanding practice: video as a medium for reflection and design," in *Design at Work: Cooperative Design of Computer Systems*, eds J. Greenbaum and M. Kyng (Hillsdale, NJ: Lawrence Erlbaum), 65–89.

Susi, T., Lindblom, J., and Ziemke, T. (2003). "Beyond the bounds of cognition," in *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum), 1134–1139.

Suthers, D. D., and Hundhausen, C. D. (2001). "Learning by constructing collaborative representations: an empirical comparison of three alternatives," in *European Perspectives on Collaborative Learning*, eds P. Dillenbourg, A. Eurelings, and Hakkarainen (Maastricht: University of Maastricht), 577–592.

Tollefsen, D. P., Dale, R., and Paxton, A. (2013). Alignment, transactive memory, and collective cognitive systems. *Rev. Philos. Psychol.* 4, 49–64. doi: 10.1007/s13164-012-0126-z

Trist, E. L. (1981). *The Evolution of Socio-Technical Systems (Occasional Paper No. 2)*. Toronto, ON: Quality of Working Life Centre.

Turner, J. R., Chen, Q., and Danks, S. (2014). Team shared cognitive constructs: a meta-analysis exploring the effects of shared cognitive constructs on team performance. *Perform. Improv. Q.* 27, 83–117. doi: 10.1002/piq.21163

Vera, A. H., and Simon, H. A. (1993). Situated action: a symbolic interpretation. *Cogn. Sci.* 17, 7–48. doi: 10.1207/s15516709cog1701_2

Wallace, D. M., and Hinsz, V. B. (2010). Teams as technology: applying theory and research to model macrocognition processes in teams. *Theor. Issues Ergon. Sci.* 11, 359–374. doi: 10.1080/14639221003729201

Wiese, C. W., Pavlas, D., and Fiore, S. M. (2011). "Towards a quantification scheme for external representations in team cognition research," in *Proceedings of 55th Annual Meeting of the Human Factors and Ergonomics Society*, (Santa Monica, CA: Human Factors and Ergonomics Society), 1437–1441.

Wildman, J. L., Salas, E., and Scott, C. P. (2014). Measuring cognition in teams a cross-domain review. *Hum. Factors* 56, 911–941. doi: 10.1177/0018720813515907

Wildman, J. L., Thayer, A. L., Pavlas, D., Salas, E., Stewart, J. E., and Howse, W. R. (2012). Team knowledge research emerging trends and critical needs. *Hum. Factors* 54, 84–111. doi: 10.1177/0018720811425365

Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322

Wiltshire, T., and Fiore, S. M. (2014). Social cognitive and affective neuroscience in human–machine systems: a roadmap for improving training, Human–Robot interaction, and team performance. *IEEE Trans. Hum. Mach. Syst.* 44, 779–787. doi: 10.1109/THMS.2014.2343996

Xiao, Y., Schenkel, S., Faraj, M. S., Mackenzie, C. F., and Moss, J. (2007). What whiteboards in a Trauma Center operating suite can teach us about emergency department communication. *Ann. Emerg. Med.* 50, 387–395. doi: 10.1016/j.annemergmed.2007.03.027

Yakura, E. K. (2002). Charting time: timelines as temporal boundary objects. *Acad. Manag. J.* 45, 956–970. doi: 10.2307/3069324

Zhang, J. (1997). The nature of external representations in problem solving. *Cogn. Sci.* 21, 179–217. doi: 10.1207/s15516709cog2102_3

Zhang, J. (1998). A distributed representation approach to group problem solving. *J. Am. Soc. Inform. Sci.* 49, 801–809. doi: 10.1002/(SICI)1097-4571(199807)49:9<801::AID-ASI5>3.0.CO;2-Q

Zhang, J., and Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cogn. Sci.* 18, 87–122. doi: 10.1207/s15516709cog1801_3

Zhang, J., and Patel, V. L. (2006). Distributed cognition, representation, and affordance. *Pragmat. Cogn.* 14, 333–341. doi: 10.1075/pc.14.2.12zha

Zhang, J., and Wang, H. (2005). The effect of external representations on numeric tasks. *Q. J. Exp. Psychol.* 58, 817–838. doi: 10.1080/02724980443000340

Zhang, Z., Hempel, P. S., Han, Y., and Tjosvold, D. (2007). Transactive memory system links: work team characteristics and performance. *J. Appl. Psychol.* 92, 1722–1730. doi: 10.1037/0021-9010.92.6.1722

CrossMark

# Macrocognition through the Multiscale Enaction Model (MEM) Lens: Identification of a Blind Spot of Macrocognition Research

Eric Laurent[1]* and Renzo Bianchi[2]

[1] Laboratory of Psychology (EA 3188), University Bourgogne Franche-Comté, Besançon, France, [2] Institute of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

> "*Given a dark room and a highly motivated subject, one has no difficulty in demonstrating Korte's Laws of phenomenal movement. Lead the subject from the dark room to the market place and then find out what it is he sees moving and under what conditions, and Korte's Laws, though still valid, describe the situation about as well as the Laws of Color Mixture describe one's feelings before an El Greco canvas.*"
>
> Bruner and Goodman (1947, p. 33)

## INTRODUCTION

Macrocognition research is concerned with cognitive processing in complex environments, goal-oriented action, goal combination and competition, cognitive-affective and cognitive-social interactions, distributed processing, and situatedness. These interests are critical to the theoretical modeling of cognitive systems for at least two reasons: (1) complexity is pervasive (and generally increases from laboratory to daily life situations), and (2) efforts are needed within (and across) all scientific fields to give meaning to, and a more global picture of, usually separate(d) knowledge fields.

In the present paper, and exactly for these two same reasons, we examine the status of "macrocognition" and suggest that *epistemologically*, "macrocognition" should not be regarded as different from other forms of cognition, including what has been called "microcognition" (Clark, 1989). Microcognition usually refers to more "internal," "subpersonal" determinants of cognitive processing (e.g., neuronal activity involved in visual perception). However, in contrast to what is sometimes found in the macrocognition literature, we do not consider microcognition as a set of "invariant" processes or "building blocks" of cognition (Letsky and Warner, 2008, p. 9). Rather, we propose here that complexity and dynamics characterize both macrocognition and microcognition. Moreover, macrocognition cannot "shunt" microcognition. Rather than promoting a new functionalism at the macroscale, we recommend that a more unitary, multiscale[1] approach to cognition be developed. Human cognition is complex and distributed, as is the biological network on which it relies. We suggest studying the generic properties of cognition through flexible analysis scales rather than creating specific fields or categories of cognition as a function of the scale of interest. In the following lines, we rely on a multiscale model of perception-action cycles' emergence, the Multiscale Enaction Model (MEM; Laurent, 2014), in which context is conceived of as being both multiple and multiscale. First, this model allows us to consider multiple interactions between processes, in line with macrocognition research's aims. Second, it highlights

---

[1] By using the term "multiscale," we refer to *multiple levels of observation and analysis* (e.g., cellular, individual, social).

the need to flexibly conceive cognitive interactions at multiple scales and to reunite cognition and aims, including basic, embodied physiological goals (e.g., hydration, energy recompletion), which do not need to be consciously elaborated.

## WHY COGNITION CAN BE "MACROCOGNITION"

The term "macrocognition" can refer to at least two perspectives over cognition. The first perspective characterizes augmented cognition theories and stresses the role of informational complexity and distributed or extended cognition. It can be opposed to more elementary views on information processing and to analytical research strategies. Macroscale factors (e.g., socioeconomical position of a family) can change microcognition (e.g., object size estimation) even in laboratory settings (Bruner and Goodman, 1947). This point is important for later discussion presented in our paper, because microcognition can neither be viewed as isolated from large-scale influences (limits of experimental-analytic approaches to cognition) nor be considered as fixed or as a set of "invariable building blocks" of cognition (limits of some macrocognition approaches, discussed later). Therefore, cognition rather appears to be enacted through interactions relating microscopic and macroscopic levels, such that a rupture between micro and macroscale analyses does not seem to be *epistemologically* sound.

The second perspective is related to the nature of the cognitive determinants that are valued, with the prefix "macro" referring to relatively large-scale influences (e.g., cognitive-social interactions), as opposed to more regional mutual influences (e.g., neuro–neuronal interactions). In this perspective "macrocognition" is often seen as being more "ecologically valid," or as enhancing "external validity" because it focuses on wide-range interactions that can be encountered in daily situations:

> "*Macrocognition is a term coined by Pietro Cacciabue and Erik Hollnagel to indicate a level of description of cognitive functions that are performed in natural (versus artificial laboratory) decision-making settings [...] the methodology for macrocognition focuses on the world outside the lab. This includes contexts designated by such terms as the 'field setting', the 'natural laboratory', and the 'real world.'"*

(Klein et al., 2003, p. 81)

We are sympathetic with the view that cognition is embedded in a network of contextual influences (i.e., the first perspective), but we anticipate limitations to the second view, which may imply a new reductionist functionalism—a large-scale equivalent to functionalist views over microcognition. Indeed, *there is no pre-set, well-suited scale of analysis*. As complexity is pervasive and multiscale, the scale at which processes should be described has to be *flexible* rather than fixed.

## WHY MACROCOGNITION CANNOT SHUNT MICROCOGNITION: FROM EXOGENOUS TO ENDOGENOUS COMPLEXITY

The term "macrocognition" usefully highlights the need for a larger scope of analysis than the one characterizing most laboratory-based experiments. However, what is usually thought of as an "external" or "environmental" factor actually combines with the organism state so that one cannot exclude any term of the interaction at any single moment. The activity of any part of an organism depends on the activity of the other parts to which it is linked. For instance, even when social-environmental complexity related to the task at hand is high (e.g., real-world lottery), factors affecting low-level biological parameters (such as ambient temperature, Cheema and Patrick, 2012) can have impacts over cognition (e.g., consumer choice). The internal resource dynamics (e.g., related to hydration) changes the willingness to make difficult gambles. Furthermore, a great amount of social psychology research, which is supposed to capture *social* complexity, is grounded in self-reported measures and individual interpretation of external complexity. In order to produce self-reports, internal construction of what is reported is a *prerequisite to* data communication and processing; this internal construction involves microscale activity (e.g., at the cellular level).

There is no macroscopic-level influence on behavior without (a) prior biological or psychological integration of the values associated with the factors of influence and (b) competition between and/or combination with the current goals and needs of the organism. Failing to recognize the complex nature of the phenomena constituting a human being can give rise to reductionism, be at the microscale or at the macroscale level. From this standpoint, suprapersonal (e.g., social), personal, and subpersonal (e.g., cellular), levels of analysis should meet. The terms "suprapersonal" and "subpersonal" refer to different scale levels in the analysis of cognition but do not imply an opposition between complexity and simplicity. Suprapersonal factors (e.g., social influences) are currently more easily detectable from a macroscale level of analysis whereas "subpersonal" factors (e.g., genetic influences) are currently more easily observable from a microscale level of analysis. However, considering one as being complex and the other one as being elementary *and* invariant would be misleading. For instance, one cannot pretend that "genetic" determinants of cognition do not involve a wealth of interacting mechanisms that influence each other (see Flint, 1999; Hill et al., 2014). In the following lines, we suggest that macrocognition and microcognition should be conceived within a single epistemological framework.

## WHY MACROCOGNITION DOES NOT *EPISTEMOLOGICALLY* DIFFER FROM MICROCOGNITION IN THE MULTISCALE ENACTION MODEL

Enactive systems produce information and knowledge by acting in their environment. In MEM (Laurent, 2014), each cell is

conceived as an autopoietic structure[2] which tends to optimize its own functioning by interacting with other cells or groups of cells. Perception-action cycles in MEM rely on those interactions because what is searched for in the environment depends on internal needs and goals. Internal needs and goals can be described at different scales. Any "external" or "ecological" influence over behavior is a *transaction* between embodied personal history (i.e., the current mode of coupling between the organism and its environment, subsequent to previous evolution and learning), goals, needs or orientations and external stimulation. Put differently, macrocognition cannot be correctly thought of without describing the interactions between the current biological state and motivation of the organism on the one hand and macroscale stimulation on the other hand.

Distributed cognition is pervasive, not only at the subpersonal level, but also at the suprapersonal level (e.g., networks of interacting individuals). There should be no epistemological rupture in the conception of distributed cognition, at physical, biological, and psychological levels. Huebner (2014) reviewed many studies suggesting that collective performance strongly depends on the coordinative properties of couples or groups, such that the collective performance cannot be reduced to the sum of individual performances. Interestingly, cognitive distribution and coordinative patterns are fundamental emerging features of groups of cells within neural networks (Craddock et al., 2013), brain areas (Bressler and Menon, 2010), and human groups (Goldstone et al., 2008). At any level, the distribution of cognition allows for the sharing of the informational load the organism is dealing with and the generation of new information through exchanges between the organism's parts. In MEM, a multiscale unifying principle is hypothesized within the central nervous system, which relates external and internal events to the organism's goals, such that both macro and microscale influences combine and are weighted as a function of their value for the organism. In MEM, the interactions between needs and goals (considered from the cellular to the psychosocial and economic levels[3]) and perception-action cycles are basic foundations for

resource allocation given the limitations in time and processing power. According to the model, teleological[4] dimensions of activity arise from the combination of need expression at the cellular and the cell network levels, and spread out to the organism and phenomenological experience through diffusion, competition, and cooperation. In this conception, the goal-directed nature of cognition makes it critical to capture any kind of influence that can modify the organism's goals. In this sense, any macroscopic-level factor should be put in the context of the organism state, as—in the other way round—the organism's informational processing and behavior should be considered in the context of larger environmental influences. In other words, in a radically distributed cognitive framework, *distribution has no pre-set scale of analysis*. Rather, distribution should be considered in every network that allows for information exchange and influences need/goal/aim satisfaction or frustration, be at cellular, cognitive, or social-affective levels. By relating micro *and* macroscale information integration to internal goals and needs, this multiscale approach provides us with tools to reunite macro and microscopic processes and levels of analysis.

## SCALE FLEXIBILITY IN DISTRIBUTED COGNITION RESEARCH: ENDING UP WITH THE BLIND SPOT OF "MACROCOGNITION RESEARCH"

"*What is a thing at one level may be relations among (different) things at another.*"

(Kelso, 1995, p. 97)

Though we subscribe to the macrocognition perspective for its emphasis on complexity, we warn the reader against the risks associated with a fixed-scale approach to cognition. Because macrocognition researchers stress the role of complexity, they should develop scale flexibility in their analyses. Even what is referred to as "macrocognition" by some researchers working on emotional context of behavior is identified as microscopic by others working on social networks. This does not change anything to the fact that, in order to analyze complex behaviors, we need to contextualize them. As a function of the scale of analysis, what can be considered a "context" varies.

Arguably what should be regarded as "ecologically valid" is the capture of multiscale interactions in experimental—or, more largely, empirical—settings that are found in everyday situations (rather than simply macroscale interactions). On those bases, and following what we discussed earlier, neglecting microscopic factors may be as harmful as neglecting macroscopic factors. In any instance of fixed-scale analysis, cognition is most probably regarded as a set of "functions" that process information under the influence of a limited number of "causes."

We invite the reader to pay attention to a blind spot that we have identified in the literature on macrocognition. The "macrocognition research blind spot" consists in associating "emergence," "dynamics" and "complexity" with macrocognition

---

[2]The Autopoiesis refers to self-production and maintenance of a "systemic variable"; an autopoietic system is a "homeostat" in which "the critical variable is *the system's own organization*" (Stafford Beer, *Preface of* Autopoiesis, *The Organization of the Living*, In Maturana and Varela, 1980, p. 66).

"*An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network.*"

(Maturana and Varela, 1980, pp. 78–79)

[3]Even if our view may be different from Maturana's regarding goals and needs, we completely agree with him when he considers that a similar organization can be found in many different structures: "any given organization may be realized through many different structures, and [...] different subsets or relations included in the structure of a given entity, may be abstracted by an observer [...] as organizations that define different classes of composite unities (Maturana, 1980, p. XX). For more information about biocomputational bases for goal and need summations, see Laurent (2014).

[4]As used here, the term "teleological" does not refer to any form of metaphysical finalism.

as opposed to "invariant processes" or "building blocks" of cognition, which would be identified by microcognition research (Klein et al., 2003). We consider this distinction misleading. As discussed earlier in this paper, microcognition is also emergent, complex, and dynamic (Laurent, 2014). Distinguishing micro from macrocognition research on the basis of emergence, complexity, and dynamics (or "reality") is neither empirically nor theoretically or logically founded. The problems associated with mainstream cognitive psychology/science (e.g., poor consideration for emergence, analytic approaches, lack of dynamic frameworks) should not be confused with the issue of the scale (i.e., micro, macro) at which the analysis is performed.

Relatedly, we do not adhere to the recurrent statements (or judgements) found in the Macrocognition literature on what "reality" is:

> "Microcognition relinquishes the coupling between the phenomenon and the real context to the advantage of the coupling with the underlying theory or model."
>
> (Cacciabue and Hollnagel, 1995, p. 57)

We rather call for a true *contextual relativism* where factors such as hydration level, laboratory settings, "internal" biological disorders, or mood fluctuations are as real as (i) the biomechanical constraints, goals, prescriptions, machines, pervasive information systems, and social context surrounding task realization and (ii) the parameters to be coordinated, which participate in emerging cognition and behaviors.

If macrocognition is to become a reference framework for the cognitive science of embedded agents, then the contexts under scrutiny should be flexibly defined, and their role theoretically reconstructed and empirically tested.

We hope that researchers interested in complexity will not add a new scale to functionalism. In other words, macrocognition should not exclude microcognition. As put by Minsky (1988), *"each higher level of description must add to our knowledge about lower levels, rather than replace it"* (p. 26). We note that this addition of knowledge should not be merely scale-specific. Rather, it should involve working on the interactions between different scales and reporting what identifies/differentiates distributed cognition at different scales. This is a basic condition to approach behavioral complexity and to develop more unitary frameworks in psychology and life sciences.

## AUTHOR CONTRIBUTIONS

EL wrote the initial draft of the manuscript. EL and RB contributed to review several versions of the manuscript and have approved the final manuscript.

## REFERENCES

Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci. (Regul. Ed).* 14, 277–290. doi: 10.1016/j.tics.2010.04.004

Bruner, J. S., and Goodman, C. C. (1947). Value and need as organizing factors in perception. *J. Abnorm. Soc. Psychol.* 42, 33–44. doi: 10.1037/h0058484

Cacciabue, P. C., and Hollnagel, E. (1995). "Simulation of cognition: applications," in *Expertise and Technology: Cognition and Human-Computer Cooperation,* eds J. M. Hoc, P. C. Cacciabue, and E. Hollnagel (Hillsdale, NJ: Erlbaum), 55–73.

Cheema, A., and Patrick, V. M. (2012). Influence of warm versus cool temperatures on consumer choice: a resource depletion account. *J. Market. Res.* 49, 984–995. doi: 10.1509/jmr.08.0205

Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing.* Cambridge, MA: MIT Press.

Craddock, R. C., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., et al. (2013). Imaging human connectomes at the macroscale. *Nat. Methods* 10, 524–539. doi: 10.1038/nmeth.2482

Flint, J. (1999). The genetic basis of cognition. *Brain* 122, 2015–2032. doi: 10.1093/brain/122.11.2015

Goldstone, R. L., Roberts, M. E., and Gureckis, T. M. (2008). Emergent processes in group behavior. *Curr. Dir. Psychol. Sci.* 17, 10–15. doi: 10.1111/j.1467-8721.2008.00539.x

Hill, W. D., Davies, G., van de Lagemaat, L. N., Christoforou, A., Marioni, R. E., Fernandes, C. P. D., et al. (2014). Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. *Transl. Psychiatry* 4:e3411. doi: 10.1038/tp.2013.114

Huebner, B. (2014). *Macrocognition.* New York, NY: Oxford University Press.

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior.* Cambridge, MA: MIT Press.

Klein, G. K., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. H., and Hollnagel, E. (2003). Macrocognition. *IEEE Intell. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Laurent, E. (2014). Multiscale Enaction Model (MEM): the case of complexity and "context-sensitivity" in vision. *Front. Psychol.* 5:1425. doi: 10.3389/fpsyg.2014.01425

Letsky, M. P., and Warner, N. W. (2008). "Macrocognition in teams," in *Macrocognition in Teams: Theories and Methodologies,* eds M. P. Letsky, N. W. Warner, S. M. Fiore, and C. A. P. Smith (Aldershot, UK: Ashgate), 1–13.

Maturana, H. R. (1980). "Introduction," in *Boston Studies in the Philosophy of Science,* Vol. 42, eds R. S. Cohen and M. W. Wartofsky (Dordrecht; Holland: D. Reidel Publishing Company), XI–XXX.

Maturana, H. R., and Varela, F. J. (1980). "Autopoiesis and cognition. The organization of the living," in *Boston Studies in the Philosophy of Science,* Vol. 42, eds R. S. Cohen and M. W. Wartofsky (Dordrecht; Holland: D. Reidel Publishing Company), 59–143.

Minsky, M. (1988). *The Society of Mind.* New York, NY: Simon and Schuster Paperbacks.

# Integrated System Design: Promoting the Capacity of Sociotechnical Systems for Adaptation through Extensions of Cognitive Work Analysis

Neelam Naikar * and Ben Elix

Defence Science and Technology Group, Melbourne, VIC, Australia

This paper proposes an approach for integrated system design, which has the intent of facilitating high levels of effectiveness in sociotechnical systems by promoting their capacity for adaptation. Building on earlier ideas and empirical observations, this approach recognizes that to create adaptive systems it is necessary to integrate the design of all of the system elements, including the interfaces, teams, training, and automation, such that workers are supported in adapting their behavior as well as their structure, or organization, in a coherent manner. Current approaches for work analysis and design are limited in regard to this fundamental objective, especially in cases when workers are confronted with unforeseen events. A suitable starting point is offered by cognitive work analysis (CWA), but while this framework can support actors in adapting their behavior, it does not necessarily accommodate adaptations in their structure. Moreover, associated design approaches generally focus on individual system elements, and those that consider multiple elements appear limited in their ability to facilitate integration, especially in the manner intended here. The proposed approach puts forward the set of possibilities for work organization in a system as the central mechanism for binding the design of its various elements, so that actors can adapt their structure as well as their behavior—in a unified fashion—to handle both familiar and novel conditions. Accordingly, this paper demonstrates how the set of possibilities for work organization in a system may be demarcated independently of the situation, through extensions of CWA, and how it may be utilized in design. This lynchpin, conceptualized in the form of a diagram of work organization possibilities (WOP), is important for preserving a system's inherent capacity for adaptation. Future research should focus on validating these concepts and establishing the feasibility of implementing them in industrial contexts.

Keywords: system design, adaptation, self-organization, sociotechnical system, cognitive work analysis

# INTRODUCTION

This paper is concerned with the design of sociotechnical systems, particularly those that are complex in nature (Vicente, 1999), such as hospitals, nuclear power plants, petrochemical refineries, military ships and aircraft, emergency management centers, and financial corporations. Designing such systems, which perform vital functions for people and society, poses considerable challenges, not least because the stakes are high—patients' lives must be saved, enemy attacks must be deterred, and natural disasters must be contained. High levels of productivity must be balanced with high levels of safety and reliability, often with shortfalls in resources, whether this is in equipment or in personnel. It is not uncommon, therefore, for these systems to operate at the edges of their effectiveness, with a fine line between successful performance and disastrous consequences. Moreover, in cases of failure, poor design has often been established as a significant contributor, with examples of such accidents including the delivery of fatal radiotherapy or chemotherapy overdoses to patients (Leveson and Turner, 1993; Institute for Safe Medication Practices, 2007), crashes of commercial airliners resulting in the deaths of hundreds of passengers and crew (Bureau of Enquiry and Analysis for civil aviation safety, 2002, 2012), military fratricide (32nd Army Air Missile Defense Command, 2003), and oil and petrochemical explosions with widespread consequences for people, infrastructure, and the natural environment (Mannan et al., 2007). Evidently, then, the question of which design philosophy and methods should underpin how these systems are conceived or formed should not be made arbitrarily.

The approach for integrated system design presented in this paper subscribes to the view that the fundamental objective in designing sociotechnical systems should be that of promoting adaptation, so that workers can deal with both routine and novel events effectively. Thus the paper begins by summarizing the empirical observations in support of this basic argument, originally formulated by Rasmussen and his colleagues (e.g., Rasmussen, 1986; Rasmussen et al., 1994). Subsequently, a case is made that designs must support actors in adapting not only their behavior but also their structure, or organization. While the importance of structural adaptation has not been unappreciated before, existing approaches for work analysis and design are limited in their capacity to support this form of adaptation. The argument is then developed, following Vicente (2002), that in designing for adaptation it is insufficient to focus on individual system elements, such as the interfaces, teams, training, or automation. Rather, the design of multiple elements must be integrated, or coordinated, such that workers are supported in adapting their structure and behavior in a coherent fashion. This paper therefore examines the capacity of current frameworks for work analysis and design to meet this objective, focusing on cognitive work analysis (CWA). Following that, the integrated system design approach is presented, which extends CWA with the intent of meeting this critical goal.

# DESIGNING FOR ADAPTATION

## Importance of Adaptation in the Workplace

A strong case has already been made that the fundamental objective in designing complex sociotechnical systems should be that of promoting successful adaptation (Rasmussen, 1986; Rasmussen et al., 1994; Vicente, 1999). This thesis, which manifests widely in one form or another (e.g., Dekker, 2003; Hollnagel et al., 2006, 2011; Hoffman and Woods, 2011; Eason, 2014; Rankin et al., 2014), is supported by a number of empirical observations.

First, complex sociotechnical systems are by and large open systems, characterized by changing or dynamic conditions (Ashby, 1956; Emery and Trist, 1965; Perrow, 1984; Gerson and Star, 1986; Rasmussen, 1986; Rasmussen et al., 1994; Vicente, 1999). This instability may result from regular perturbations, either within the system (e.g., technical malfunctions, staffing shortages) or in the external environment (e.g., economic fluctuations, changing weather patterns). Moreover, these systems may have to contend with novel circumstances, or events that cannot be fully predicted a priori, such as a new kind of military threat (Reich et al., 2010; Herzog, 2011), an unexpected reaction of a patient to an anesthetic during surgery (Hoppe and Popham, 2007), or an unforeseen chain of supplier collapses in the wake of a natural disaster (Park et al., 2013). These systems, therefore, must be capable of continuously and reliably dealing with significant variability in their work environments.

Studies of complex sociotechnical systems have also demonstrated that the greatest threats to these systems' effectiveness are posed by unanticipated events (e.g., Rasmussen, 1968a,b, 1969; Perrow, 1984; Reason, 1990; Leveson, 1995; Vicente, 1999). As these situations cannot be predicted, analysts or designers cannot provide workers with "ready-made" solutions for handling these events. Moreover, as these situations are unfamiliar to workers, they cannot simply retrieve a suitable solution from their portfolios of prior experiences. Instead, workers must respond flexibly and creatively to deal with these situations successfully (e.g., Rochlin et al., 1987; Bigley and Roberts, 2001; Bogdanovic et al., 2015) and thus finish the design (Rasmussen and Goodstein, 1987).

Aside from dealing with unexpected events, adaptations are necessary regularly, or even routinely, in everyday situations (Simon, 1969; Gerson and Star, 1986; Rasmussen, 1986; Suchman, 1987; Weick, 1993; Rasmussen et al., 1994; Vicente, 1999). Even small changes in context may require adaptation (Vicente, 1999), and it is not possible to formulate an algorithm, plan, or procedure for every single complication (Hoffman and Woods, 2011), even if it were safe to do so (Dekker, 2003). Thus everyday work requires ongoing local adjustments or improvisations to accommodate the inevitable flux that arises in the system (Bigley and Roberts, 2001; Rankin et al., 2014; Bogdanovic et al., 2015; Militello et al., 2015).

Another significant observation is that adaptations are important not just for safety but also for organizational productivity and workers' health (Vicente, 1999). In computerized workplaces, where routine tasks are typically

automated, system success can hinge on the capacity of workers to conjure up innovative solutions to emerging problems for which algorithms have not been, or cannot be, written. Furthermore, it has long been recognized that workers with greater decision latitude tend to have better health, as indicated by such factors as longevity and the absence of stress or disease (Karasek and Theorell, 1990; Vicente, 1999; Eason, 2014). Such workers have the autonomy to decide how to manage their work demands, including the ability to improvise or adapt in doing their jobs, and to follow their individual preferences when it is appropriate to do so.

Finally, while the importance of adaptation in the workplace is clear, it is also evident that ongoing adaptation to changing situations and unforeseen circumstances can be demanding (Rasmussen, 1986; Rasmussen et al., 1994; Vicente, 1999; Dekker, 2003; Hoffman and Woods, 2011; Bogdanovic et al., 2015). The context or conditions under which adaptation is required, as it is experienced by workers, is usually exacting, involving multiple, conflicting goals, significant time pressure, many unexpected turns of events, and considerable stress stemming from the awareness of the potentially disastrous consequences of failure. Furthermore, adaptation can be an intellectually or cognitively challenging exercise, involving very complex reasoning under demanding conditions (Rasmussen et al., 1994; Dörner, 1996; Vicente, 1999). Typically, workers must make rapid decisions about whether, when, and how to adapt in light of their judgments of the local conditions, awareness of the broader organizational goals and constraints, and assessments of the risks and opportunities this context presents (Dekker, 2003).

Workers, therefore, should not have to—or be expected to—adapt in an *ad hoc* manner, using technology or workplace designs that do not support or, worse still, deliberately inhibit improvisation, as is so often the case (Vicente, 1999; Eason, 2014). Aside from placing, quite unnecessarily and unfairly, an increased burden on workers who are already working under very demanding conditions, this situation could lead or contribute to unsafe or unproductive outcomes. Instead, workers should be provided with systematic support through the system design, including the design of technology, training, and procedures, to help them in adapting seamlessly and successfully to the unexpected and changing demands of their jobs (Rasmussen, 1986; Rasmussen et al., 1994; Vicente, 1999; Dekker, 2003; Eason, 2014; Rankin et al., 2014; Militello et al., 2015).

## Behavioral and Structural Adaptation

If we are to design systems that facilitate successful adaptation, a key question that arises is what manner of adaptations are needed in the workplace, and thus should be deliberately supported through design. The following studies demonstrate the importance of both behavioral and structural adaptation to system effectiveness. Greater emphasis is placed on illustrating the nature of structural adaptation in the workplace, since existing analysis and design approaches are limited in supporting this form of adaptation, as discussed in more detail later in this paper.

Empirical studies of workers in complex sociotechnical systems reveal that one form of adaptation that occurs entails

actors adapting their behavior, or effectively adjusting their tasks, plans, goals, actions, or priorities in step with the unfolding situation. Bigley and Roberts (2001) provide a detailed account of the improvisations they observed during a field study of a large fire department employing the incident command system, a widespread approach for emergency management in the United States of America. They categorized the improvisations as involving tools, rules, and routines. When a truck arrives at the scene of an emergency, for instance, personnel may have no choice but to improvise with the tools available on the truck, employing them in unusual ways to handle the situation. In other cases, the adaptations may include departures from rules, directly breaching standard operating procedures. As an example, one procedure prohibits firefighting teams from approaching a fire from opposite positions, as one group can push the fire into another. However, a firefighter discussed a situation in which "opposing hose streams" was in fact used as the primary tactic. Lastly, the execution of standard routines, such as those for "hose laying" or "ladder throwing," may also be adjusted to accommodate local contingencies. According to Bigley and Roberts, such improvisations are regarded as legitimate within the organization, provided they are consistent with organizational goals and are unlikely to harm personnel or other people.

Observations of behavioral adaptation in the workplace have also been documented in a number of other contexts. Goteman and Dekker (2001), for example, discuss how commercial pilots shed tasks when confronted with demanding circumstances, postponing some jobs until the situation becomes more manageable. Similarly, Militello et al. (2015) observed that military pararescue teams are constantly juggling priorities for evacuating injured personnel from hostile areas, depending on what transpires at the scene in relation to such factors as the urgency of patients' medical conditions, the actions of adversaries, and the available resources. Finally, within a health care context, Bogdanovic et al. (2015) discuss how surgeons may interrupt a surgical procedure on discovering unanticipated patient states, such as the presence of inflammation, in order to discuss the next steps with the medical team.

Further to such adaptations in workers' activities, empirical studies provide considerable evidence for structural adaptation, whereby multiple actors are involved in adjusting their structure or organization in line with the emerging situation. As a result, the particular actors involved and their roles and relationships may be constantly changing. A potent example is provided by Rochlin et al. (1987), who conducted a field study of how navy personnel on aircraft carriers coordinate their work activities. Rochlin et al. found that the formal organization of this system—that which is documented on paper—is rigid, hierarchical, and centralized, being characterized by clearly defined chains of command and means to enforce authority. Typically, this organizational structure governs operations on the ship.

During complex operations, however, Rochlin et al. (1987) found that a very different type of organizational structure is adopted. This organizational structure may be described as informal, given that it is not officially documented. The informal organization is flat and distributed rather than hierarchical and

centralized. For instance, based on their access to information, lower-ranked personnel have the autonomy to make critical decisions without the approval of officials with higher rankings, especially when faced with significant time constraints. The informal organization is also flexible in that there is no pre-specified plan for when it will be adopted. Moreover, the specific organizational structure that is adopted on any one occasion is emergent, such that there is no simple or fixed mapping between people and roles and therefore no single informal organization. Instead, the work organization on the ship adapts to changes in circumstances. According to Rochlin et al. this adaptability contributes greatly to balancing the need for safety with the push for productivity.

Bigley and Roberts's (2001) observations of a fire department employing the incident command system for emergency management echo many of Rochlin et al.'s (1987) findings. At one level, this system is highly formalized with an extensive set of policies, procedures, and instructions. Jobs are specialized and have very particular training requirements. In addition, positions within the system are arranged hierarchically and reflect formal authority relationships. Objectives and plans are established near the top of the hierarchy and serve as a basis for guiding decisions and behaviors at lower levels. Nevertheless, as Bigley and Roberts discovered, the fire department consistently employs a number of mechanisms for rapidly converting this rigid organizational structure into highly flexible arrangements suitable for dealing with the specific emergencies encountered. Bigley and Roberts describe these mechanisms as involving structure elaborating, role switching, authority migration, and system resetting.

Structure elaborating describes the process of organization construction at the scene of an incident, with the first captain arriving becoming the incident commander, at least temporarily. After assessing the situation and developing an initial plan, the incident commander begins to build an organization by assigning roles and tasks to incoming resources, a process which may continue until the emergency shows signs of subsiding. Pre-existing roles or positions within the incident command system are filled with people only to the extent required, perhaps with more positions becoming filled as the situation unfolds. Furthermore, some functions may not be assigned to specialized positions until it is necessary to do so, with personnel already established in particular positions being responsible for multiple functions in the meantime.

Role switching sums up the observation that positions continue to be activated and relationships established in line with the emerging situation. In addition, positions are deactivated when the appropriate role structure for an emergency changes, and personnel are either shifted into different positions or discharged. Authority migration recognizes that although formal authority relationships remain fixed, informal decision-making authority can migrate rapidly to personnel possessing the most relevant expertise. Thus senior personnel may defer to lower-level experts who are more technically qualified given the specific characteristics of the emergency, temporarily shifting authority to them. Lastly, system resetting involves disengaging or regrouping. When the current approach appears to be having no effect or is found to be unsuitable because of unexpected

occurrences, the team is withdrawn from the situation and reconfigured or redirected. As Bigley and Roberts observe, "Within the most reliable systems, objectives and corresponding structural elements and relationships are adjusted swiftly in accordance with changing environmental contingencies" (p. 1287).

Finally, Bogdanovic et al. (2015) provide a detailed account of how the task distribution among actors in surgical teams alters as a function of specific occurrences during surgery. According to Bogdanovic et al. only the general task distribution is established prior to the surgical procedure. While the delegation of some tasks are determined by team members' professions, such as whether one is an anesthetist, nurse, or surgeon, tasks that can be fulfilled by any person are not assigned in advance but are delegated dynamically throughout the surgery, depending on the circumstances. Some options for the task distribution in view of the anticipated challenges may be contemplated before surgery. However, if unforeseen complications arise, new arrangements are conceived and instituted at the time.

A specific reason tasks may be redistributed during surgery is that problems emerge for which a team member does not possess the necessary skills. Thus a senior physician may take over a step of the procedure initially assigned to someone else. Another possibility is that the procedure itself may need to be altered because of the specific problems encountered, such that the steps of the revised procedure must be reassigned among team members. Team members will also assist their colleagues to balance the workload within the group. An anesthetist, for example, may help the scrub nurse if the circulating nurse is busy. Lastly, the task distribution may change as a result of additional resources being mobilized for the task at hand. For instance, due to unforeseen complications during surgery, it may be necessary to call a more experienced clinician for help. According to Bogdanovic et al. (2015), such open-ended fine tuning of the task distribution, including the temporary assistance provided by team members across their professional demarcations, provides the flexibility necessary for dealing with situational variability, minimizes pressure on the team, and enables a smoothly running procedure.

## Necessity of Integrated System Design

The preceding discussion has clear implications for system design. First, designing for adaptation is essential so that workers can handle a wide variety of events, including both routine and novel ones, effectively. Moreover, workers must be supported in adapting both their behavior and structure, effortlessly and seamlessly. It is important to recognize that changes in behavior may or may not be associated with changes in structure. In addition, changes in structure may be associated with behavioral opportunities not available to workers otherwise. Irrespective of these fine distinctions, designing for adaptation must encompass the behavioral and structural possibilities comprehensively if we are to create systems that are resilient in the face of instability and uncertainty.

Evidently, systems are comprised of multiple elements, which must work together in concert in view of a common purpose. Consequently, the aforementioned objectives cannot be achieved

by focusing on the design of individual elements, such as the interfaces, teams, training, or automation. In the context of promoting worker adaptation, the need for integrated system design was emphasized by Vicente (2002). He observes that designing for adaptation cannot be achieved in a piecemeal fashion. That is, a system will not necessarily be adaptive simply because it has an ecological interface, even though such interfaces are intended to support adaptation (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1990, 1992). Instead, to create systems that can adapt successfully, all of the different elements must be designed in a coordinated manner based on a common philosophy, specifically a philosophy focused on promoting adaptation. Naikar (2012) echoes these observations, recognizing in particular that a system will not necessarily be adaptive solely on the basis of its team design, even if that is intended to engender flexibility (Naikar et al., 2003). In this paper, we elaborate on these ideas by taking into account the empirical observations described above.

To create adaptive systems, the design of multiple elements must be integrated based on a common philosophy that promotes both structural and behavioral adaptation. It is also clear that to preserve a system's inherent capacity for adaptation to novelty, the designs of the different elements must support the full range of opportunities for structural and behavioral adaptation in the workplace and that they must do so uniformly across multiple actors in the system. Thus, if a team design supports possibilities for structural or behavioral adaptation that an interface design does not, the design of the two elements would not be integrated, or compatible, with respect to the goal of promoting adaptation. Similarly, if an interface design for an actor or group of actors in a system supports possibilities for adaptation that are not recognized or accommodated by the interface designs for other actors in the system, such that some or all of the possibilities cannot be realized by any of the actors, the design of this element would not be integrated across multiple actors in the system. Such approaches would not necessarily foster successful performance in the event of change or novelty, and they might even inhibit it. Moreover, as demonstrated later, simply approaching the design of multiple elements concurrently with the philosophy of promoting worker adaptation may be insufficient to achieve this level of integration. Rather, the design framework must encompass explicit mechanisms for binding or anchoring the designs of multiple elements, so that the system design supports the range of possibilities for adaptation in structure and behavior, across multiple actors, in a coherent fashion.

## WORK ANALYSIS AND DESIGN

Designing for adaptation requires special approaches for work analysis, as the way in which the work demands of a system are understood is tightly integrated with how those work demands are supported through design. As is well established now, work analysis techniques may be differentiated on the basis of whether they are normative, descriptive, or formative in orientation (Rasmussen, 1997; Vicente, 1999). The following discussion demonstrates briefly that normative approaches are unsuitable

for designing for adaptation, whereas descriptive approaches are insufficient. Instead, a formative approach is necessary.

Normative approaches, such as task analysis techniques that define sequences or timelines of tasks (Kirwan and Ainsworth, 1992), are concerned with specifying the ideal ways in which to perform work under particular conditions. However, in open systems, which are subject to situational variability, the anticipated conditions may never match the conditions that are experienced precisely, such that the recommended task sequences or procedures may not in fact be the most productive or safest way of handling the situation. Moreover, removing autonomy from workers in deciding the best way of performing a task or in following their individual preferences when it is appropriate to do so may be counterproductive for workers' health and ultimately for organizational productivity.

Descriptive approaches, such as some of those described in Schraagen et al. (2000), are concerned with developing a faithful understanding of the cognitive challenges that workers experience in their jobs and the cognitive strategies they employ for dealing with these challenges. On this basis, designs can be developed that support workers in handling these challenges more effectively and that accommodate the variability in work practices observed in everyday work. One limitation of such approaches, however, is that the resulting appreciation of cognitive challenges and viable cognitive strategies is generally constrained to familiar, recurring, or anticipated conditions, which can be studied or observed. The capacity of such approaches to support adaptation to unforeseen events, then, is limited to the extent to which the existing challenges and strategies are relevant to the novel conditions. Descriptive techniques, therefore, must be complemented with a formative approach to work analysis, and CWA offers a suitable starting point.
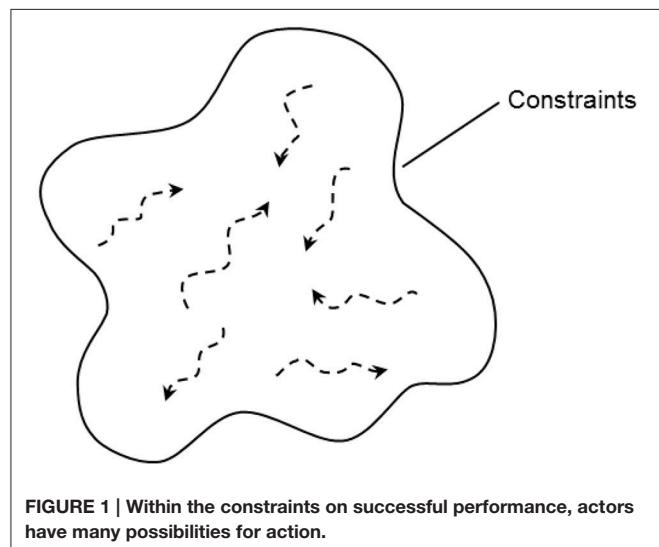
## Cognitive Work Analysis

CWA is a comprehensive framework for modeling the work demands on actors in terms of the constraints, or boundaries, that must be upheld by their actions irrespective of the particular conditions they are faced with (Rasmussen, 1986; Rasmussen et al., 1994; Vicente, 1999). Thus this framework is concerned with the constraints that are applicable not only in familiar, recurring, and anticipated situations but also in situations that cannot be predicted a priori. Although these constraints must be observed or respected for effective performance, such that they bound the possibilities for action available to actors, within these constraints actors still have many degrees of freedom for action, as indicated by the trajectories in **Figure 1**. Therefore, using this framework, designs can be developed that deliberately provide actors with the flexibility to adapt their work practices to a wide range of situations without crossing the boundaries of successful performance. In contrast to normative and descriptive approaches, then, which focus on specifying how work *should be done ideally* or *is done currently* in a system, CWA is a formative approach that is concerned with specifying the constraints that bound how work *can be done effectively*.

The CWA framework comprises five dimensions, which are concerned with different types of constraints (**Table 1**). These

| Dimensions | Constraints | Modeling Tools |
|---|---|---|
| Work domain analysis | Work domain—constraints placed on actors by the physical, social, and cultural environment, including the system's purposes, values and priorities, functions, and physical resources | Abstraction-decomposition space, abstraction hierarchy (Rasmussen et al., 1994; Vicente, 1999; Naikar, 2013) |
| Activity analysis | Activity—constraints placed on actors by the activities necessary in the system to achieve the system's purposes, values and priorities, and functions with the available resources | Contextual activity template (Naikar et al., 2006), decision ladder (Rasmussen et al., 1994; Vicente, 1999) |
| Strategies analysis | Strategies—constraints placed on actors by the cognitive strategies that can be utilized for achieving the activities necessary in the system | Information flow map (Rasmussen et al., 1994; Vicente, 1999) |
| Social organization and cooperation analysis | Work organization—constraints placed on actors by the ways in which work can be allocated, distributed, and coordinated in the system | Diagram of work organization possibilities |
| Worker competencies analysis | Workers—constraints placed on actors by the ways in which the work demands of the system can be met given human cognitive capabilities and limitations | Skills, rules, and knowledge taxonomy (Rasmussen et al., 1994; Vicente, 1999) |



**FIGURE 1 | Within the constraints on successful performance, actors have many possibilities for action.**

dimensions collectively define a constraint-based space, such as that illustrated in **Figure 1**, in relation to the system of interest. As shown in **Table 1**, each CWA dimension has special modeling tools for capturing and representing the various constraints on actors. In the current CWA framework, the social organization and cooperation dimension takes advantage of the modeling tools from the preceding dimensions (Rasmussen et al., 1994; Vicente, 1999). However, in this paper the diagram of work organization possibilities (WOP) is introduced as a special modeling tool for this analysis.

## Value of Cognitive Work Analysis for Design

Considerable empirical evidence exists for the value of CWA for design, specifically in relation to ecological interface design, a framework that utilizes CWA as a basis for designing interfaces for workers in complex sociotechnical systems (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1990, 1992). For example, as documented in existing reviews (Vicente, 2002; Naikar, 2012), controlled experiments have demonstrated the value of
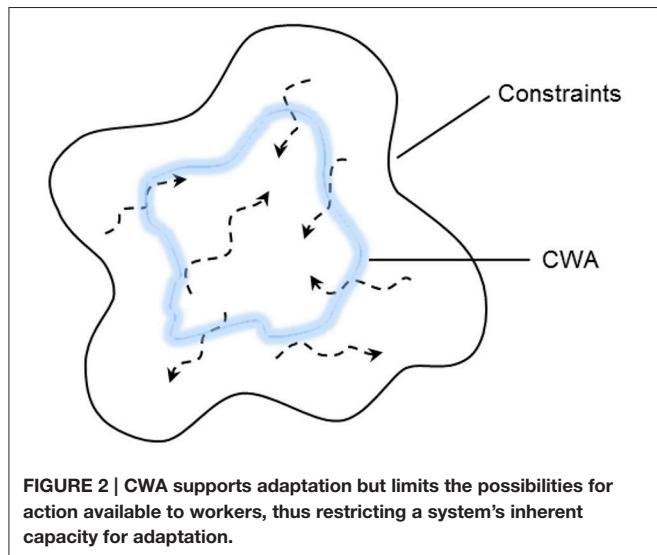
ecological interface design for process control (Christoffersen et al., 1996; Pawlak and Vicente, 1996; Reising and Sanderson, 1998, 2000a,b; Ham and Yoon, 2001; Jamieson, 2007; Lau et al., 2008), information retrieval (Xu et al., 1999), neonatal intensive care (Sharp and Helmicki, 1998), network management (Burns et al., 2003), aviation (Borst et al., 2006), and military command, and control (Bennett et al., 2008). Collectively, the results of these studies demonstrate that ecological interface design can be applied to a range of systems and that, for those systems, this framework can uncover novel information requirements that can lead to better performance by workers in comparison with that obtained with existing interfaces.

The value of CWA for problems other than interface design has also been demonstrated. Detailed industrial case studies have shown, for example, that CWA can be used for selecting system designs (Naikar and Sanderson, 2001), designing teams (Naikar et al., 2003), and developing training systems (Naikar and Sanderson, 1999) that promote flexibility. As these applications of CWA were executed in industrial settings, experimental investigations were unfeasible. However, the value of CWA for these applications was demonstrated on the basis of its ability to impact practice, its uniqueness in comparison with the design outcomes obtainable with conventional approaches, and its feasibility of implementation within a project's schedule, personnel, and financial resources (Naikar, 2013). These criteria are more commonly applied for assessing worth in industrial practice (Whitefield et al., 1991; Czaja, 1997; Vicente, 1999).

## Limitations of Cognitive Work Analysis for Design
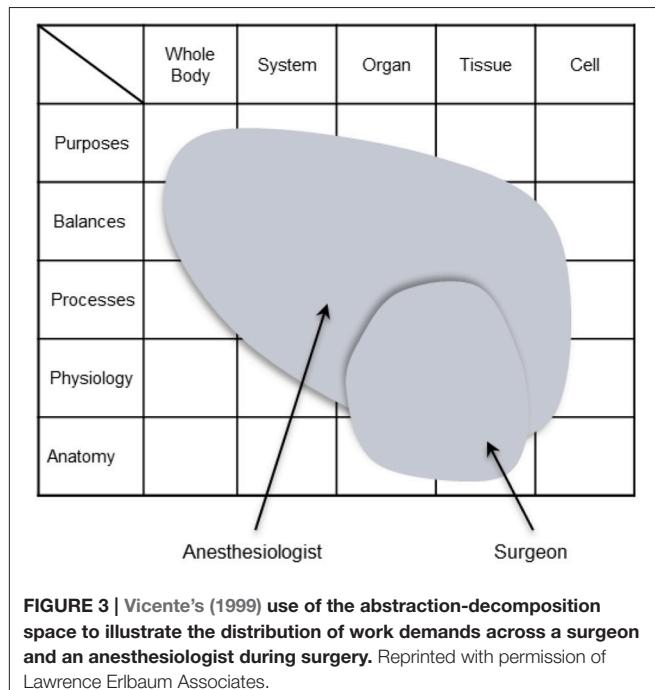
While it is clear that CWA can support adaptation, in this paper we observe that this framework has two, related, limitations that could restrict a system's inherent capacity for adaptation (**Figure 2**). The first has to do with the capacity of this framework to support adaptations in the work organization, or structural adaptation. The second concerns its capacity to facilitate the integration of multiple system elements to produce an integrated system design.

One reason that CWA is limited in its capacity to promote adaptation is that although this framework can support actors

**FIGURE 2 | CWA supports adaptation but limits the possibilities for action available to workers, thus restricting a system's inherent capacity for adaptation.**



**FIGURE 3 | Vicente's (1999)** use of the abstraction-decomposition space to illustrate the distribution of work demands across a surgeon and an anesthesiologist during surgery. Reprinted with permission of Lawrence Erlbaum Associates.

in adapting their behavior, in its current form it does not necessarily support actors in adapting their structure, especially in unforeseen situations. Yet, as the empirical studies described earlier in this paper and elsewhere show, adaptations in the work organization are also critical for successful performance. The fundamental texts on CWA by Rasmussen et al. (1994) and Vicente (1999) do recognize that complex sociotechnical systems are characterized by flexible organizational structures, such that the structures actors adopt may vary subtly or significantly in response to the local context. Thus they point out that the social organization and cooperation dimension of CWA must be concerned with the various organizational structures that are relevant. Moreover, the texts observe that shifts in structure are governed by such criteria as the competencies of actors, the access actors have to information or the means for action, the requirements for safety and reliability, the need for compliance with policies and regulations, the requirements for workload sharing, and the need for minimizing coordination demands. However, neither text offers a formative approach for analyzing the work organization. Instead, the suggested approach seems descriptive in orientation as it appears to be concerned with organizational structures that can be observed or are judged to be reasonable in recurring classes of situation (Naikar and Elix, 2016a).

As a case in point, Vicente (1999) discusses that, within the CWA framework, the analysis of organizational structures is undertaken in relation to particular classes of situation and, to illustrate this approach, he provides an example of how CWA can be used to analyze the organizational structures in a health care system. Specifically, he describes how the work demands of surgery may be distributed differently across a surgeon and an anesthesiologist, and he points out that the distributions of work demands may change if the patient is in pre-operation rather than in surgery. Furthermore, to complement his discussion, he illustrates how models from the CWA framework may be used for representing such distributions (**Figure 3**). However, in this approach, CWA is being used to describe the organizational

structures that are adopted by workers in recurring classes of situation, rather than to understand the structures that can be adopted irrespective of the situation. This approach may be useful for developing designs that support workers in commonly occurring situations, which is important. However, designs based on this approach may not be suitable for dealing with some kinds of situational variability or with unanticipated events particularly, because they may not support the organizational structures that are relevant—or that emerge—in unforeseen circumstances. Moreover, as these structures may present new behavioral opportunities, the resulting designs may not support some behavioral possibilities.

Another, related, reason that CWA is limited in its capacity to facilitate adaptation concerns its ability to support integrated system design, whereby the design of multiple elements are coordinated across multiple actors in the system, such that workers are supported in adopting the range of possibilities for structural and behavioral adaptation in a unified manner. As discussed in more detail in the next section, to facilitate the integration of multiple elements in a way that promotes adaptation, the design of each element must be anchored to a common set of constraints. In complex sociotechnical systems, which are comprised of multiple actors, the full set of constraints that is relevant to each actor, or group of actors, in the system is dependent on the organizational structures that are possible. Accordingly, the design of each element must be coordinated around the organizational constraints. Hence the lack of a formative means for analyzing the organizational structures that are relevant, irrespective of the situation, does not limit simply the capacity of CWA to promote structural adaptation but also its capacity to facilitate the integration of multiple elements, across multiple actors, to produce an integrated system design.

We do not suggest here that a formative analysis of the work organization is sufficient for creating an integrated system design. It is also important, for example, to have systematic processes for respecting the organizational constraints in the design of each element, as discussed in more depth later. The formative analysis of organizational structures, however, is a central step in creating an integrated system design. Perhaps it is also worth making the point explicitly that a formative analysis of the work organization in itself does not guarantee that multiple elements will be considered in the design process, but, once again, this analysis is essential for the designs of multiple elements to be well integrated, as elaborated in the next section.

Finally, it is worth noting that existing design approaches based on CWA are limited in their capacity to promote adaptation in the manner concerned with here. First, detailed design approaches are focused largely on individual system elements, such as the interfaces (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1990, 1992) or teams (Naikar et al., 2003; Naikar, 2013), although this is not to say that the need for integration with other elements was unappreciated. In relation to system design, Vicente (1999) makes the observation that particular phases of CWA can be used to inform particular classes of system design interventions. For example, he discusses that work domain analysis can be used to inform the design of information systems, that social organization and cooperation analysis can be used to inform the design of teams, and that worker competencies analysis can be used to inform the design of training programs. However, it is unclear how Vicente (1999, 2002) intended the designs of the different elements to be integrated (Naikar and Elix, 2016a). If the designs of these elements are informed by different phases of CWA, such that they are based on distinct sets of constraints, the resulting designs would not necessarily support the same possibilities for adaptation. Alternatively, if the design of each element is based on all five phases of CWA, the resulting designs may be integrated but only in relation to a reduced space of possibilities for action, as the analysis would be restricted deliberately to organizational structures that can be observed or are judged to be reasonable in recurring classes of situation.

Further to Vicente (1999, 2002), some approaches have addressed how particular phases of CWA can be used to support different stages of the system lifecycle, such as requirements definition, design, and evaluation, and to support the design of a variety of system elements, such as the interfaces, teams, and training (Sanderson et al., 1999; Hori et al., 2001; Read et al., 2015a,b). It would be fair to say that all of these approaches recognize at some level the need for the design of multiple elements to be integrated in some fashion, although Hori et al. (2001) and Sanderson et al. (1999) do not address this point explicitly. Read et al. (2015b) discuss the need to ensure that the design of all of the elements are coordinated and, in the context of a case study, Read et al. (2015a) describe the use of a template for summarizing a design concept, which requires that design features associated with all system elements are documented. On the basis of the information provided in these papers, it seems that this process could help to ensure that the designs of multiple elements are considered concurrently, although from

the case study it appears that this is not a guaranteed result, given the ratings of the four participants in the design process and the analyst's reflections. In any case, assuming all elements are considered concurrently, it is unclear in what way, or on what basis, the design of the different elements would be coordinated using the process described, and thus what manner of integration the process would promote. However, considering that the process is based on the existing CWA framework, one can assume that it would be limited in its capacity to support structural adaptation and to facilitate the integration of multiple system elements in the fashion with which this paper is concerned.

## INTEGRATED SYSTEM DESIGN

This paper proposes an approach for integrated system design, based on extensions of CWA. The approach develops substantially ideas described initially by Naikar (2006, 2012, 2013) for the analysis of the work organization and by Naikar and Elix (2015) for coordinating the design of multiple system elements. The express intent of this approach is to promote the capacity of sociotechnical systems for adaptation.

The proposed approach has two particular premises. First, the approach presupposes that complex sociotechnical systems are comprised of multiple actors, as a single actor could not possibly attend to all of a system's work demands (**Figure 4**). For example, a single actor could not possess or develop the full set of knowledge and skills necessary for dealing with all of the system's work demands effectively. Similarly, a single actor could not have the physical and mental capacity to cope with all of the system's work demands in the combinations and pace at which they occur. The significance of this straightforward assumption is made clear later.

Another premise of the proposed approach is that in complex sociotechnical systems there is usually no single or best way of organizing work, or of distributing the work demands across multiple actors. Instead, as empirical studies such as those cited earlier (Rochlin et al., 1987; Bigley and Roberts, 2001; Bogdanovic et al., 2015) show, flexible work structures that can be adapted to local contingencies are necessary for dealing with the demands of a range of situations, including unforeseen events. This means, then, that designs must support actors in adapting not only their behavior but also their structure, such that it is possible for actors to meet the demands of a variety of circumstances, some of which may be completely novel to them.

In line with these premises, the proposed approach for integrated system design recognizes that to promote the capacity of sociotechnical systems for adaptation, it is necessary to understand the set of possibilities for work organization in a system irrespective of the situation. From a design perspective, this is necessary not simply for supporting multiple actors in adapting their structure but for coordinating the design of multiple elements, such as the interfaces, teams, training, and automation. As a result, actors will be supported in adapting their structure as well as their behavior —in a unified fashion—to meet the demands of a range of circumstances. Accordingly, the approach places emphasis on demarcating the set of possibilities

for work organization in a system, given the system's constraints, and subsequently developing designs for each element that can accommodate the range of possibilities. These ideas are elaborated in the following discussion.

For the purposes of integrated system design, the set of possibilities for work organization in a system is delineated through extensions of CWA, rather than any other work analysis technique, as a formative approach is necessary for supporting adaptations in both behavior and structure across a range of situations. As demonstrated in detail later, the possibilities can be delineated within the social organization and cooperation dimension of CWA (**Table 1**) by applying the criteria that govern shifts in work organization in a formative manner to examine how the work demands of the system can be distributed across actors—both human and automata. Ideally, the work demands would be derived from the first three dimensions of CWA, namely work domain analysis, activity analysis, and strategies analysis. However, given practical considerations, the work demands may be derived solely from work domain analysis, as it encompasses both novel and anticipated situations (Naikar and Elix, 2015, 2016a). Once the organizational possibilities have been defined, designs for each of the system elements can be developed to support those possibilities at the three levels of cognitive control that actors can bring to the performance of a task. These three levels of cognitive control, skill-based, rule-based, and knowledge-based behavior, are considered within the worker competencies dimension of CWA. Thus the proposed approach coordinates the design of multiple system elements around the organizational constraints.

The set of possibilities for work organization in a system is regarded as the central mechanism for integrating the design of multiple elements because complex sociotechnical systems are comprised of multiple actors. To create an integrated system design, one in which all of the elements support adaptation in a coherent fashion across multiple actors, the design of each

element must be anchored to a common set of constraints. Given multiple actors, the constraints of the work domain, activity, strategies, and workers that are applicable to an actor, or group of actors, are dependent on the possibilities for work organization (**Figure 5**). Hence the design of each element, for each actor, must be coordinated around these possibilities, or organizational constraints. While the design of each element must also respect the constraints of the work domain, activity, strategies, and workers, the designs of these elements can only be coordinated around those constraints if it is assumed that a single actor is responsible for all of the system's work demands. However,
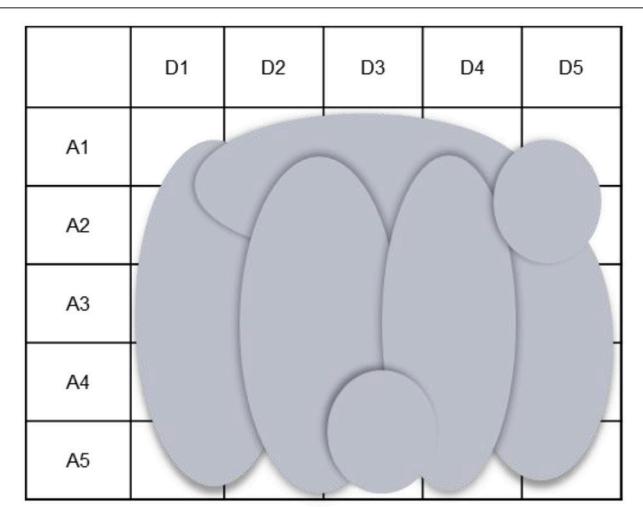


**FIGURE 5 | Use of the abstraction-decomposition space to illustrate that when there are multiple actors, the constraints that are relevant to an actor, or group of actors, are dependent on the possibilities for work organization.** "A" signifies a level of abstraction whereas "D" signifies a level of decomposition.
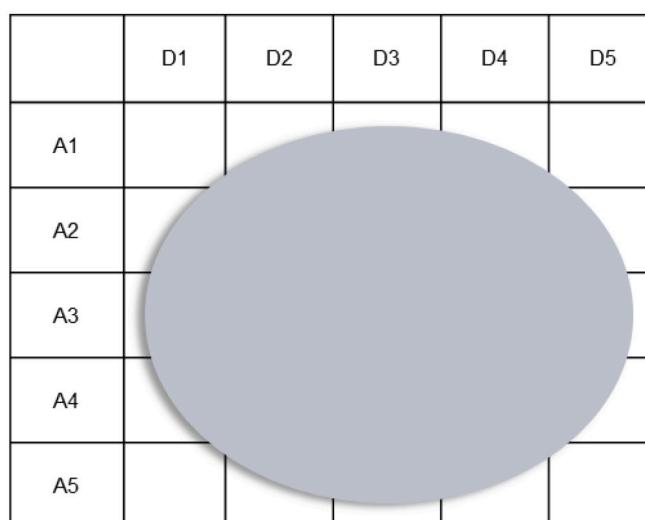


**FIGURE 4 | Use of the abstraction-decomposition space to emphasize that a single actor could not possibly attend to all of a system's work demands.** "A" signifies a level of abstraction and "D" signifies a level of decomposition.

this design approach is unsuitable for complex sociotechnical systems, as multiple actors are necessary for fulfilling the system's work demands.

Notably, as the constraints that are relevant to a particular actor or group of actors are dependent on the possibilities for work organization, understanding the set of possibilities is essential not only for supporting actors in adapting their structure but also in adapting their behavior. As indicated earlier, the different structural possibilities are associated with distinct behavioral opportunities. Therefore, to appreciate the full set of behavioral possibilities available to particular actors, it is necessary to establish the full set of work structures in which they can participate. Otherwise, the resulting constraint-based space for each actor will be smaller than their actual space of possibilities for action. This means that the associated designs, though offering some degree of flexibility to each actor, will limit the possibilities for action available to them, ultimately restricting the capacity of the system for adaptation.

By emphasizing the necessity of defining the set of possibilities for work organization independently of the situation, the proposed approach promotes greater adaptation than can be achieved by focusing designs on a subset of possibilities. For example, the approach can lead to designs that support greater adaptation than designs based on work structures observed in recurring situations. Similarly, it can lead to designs that promote greater adaptation than those based on work structures deemed ideal under certain conditions. This approach, then, can foster the development of more robust or resilient systems that are capable of coping with idiosyncratic circumstances or situations involving small variations from recurring or pre-defined conditions, as even small changes in context can require adaptation by workers. Moreover, it can foster the development of systems with greater capacity to deal with novel events, which is particularly important given that these events are widely regarded as posing the most significant threats to performance (Rasmussen, 1968a,b, 1969, 1986; Perrow, 1984; Reason, 1990; Rasmussen et al., 1994; Leveson, 1995; Vicente, 1999).

The proposed approach therefore enhances the quality of the integration of multiple system elements, with respect to the goal of promoting adaptation, compared with that achievable by designing the various elements using existing design approaches based on CWA. As an illustration, the application of existing approaches to design particular elements could involve using the ecological interface design framework (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1990, 1992) to create the displays for a system and a technique described by Naikar et al. (2003; also see Naikar, 2013) to develop the team designs for that system. However, applying these techniques in combination would not necessarily ensure that the designs of the two elements are well coordinated, particularly because there is no explicit mechanism for tying together, or binding, the designs of the interfaces and teams across multiple actors in the system.

In particular, the ecological interface design framework cited above is based on the constraints of the work domain and workers, whereas the team design approach is concerned with the constraints of the work domain and activity. Notably, Bennett and Flach (2011) describe an approach for ecological interface design that incorporates the constraints of the work domain, activity, and workers. Nevertheless, even if the designs of both elements were anchored somehow to a common set of constraints, whether this is the constraints of the work domain, activity, workers, or all of these constraints, this approach would be insufficient for complex sociotechnical systems.

Assuming that the existing techniques for both elements involve some kind of recognition, formal or informal, of there being multiple actors and of there being different ways of organizing work among these actors, as the team design technique does at least, the resulting designs would most probably take into account only a subset of the work organization possibilities, say those that can be observed or anticipated. Consequently, while the designs of the two elements may be integrated across multiple actors in the system, by anchoring the designs of both elements to the constraints considered relevant to each actor or group of actors, the designs would be integrated only in relation to a reduced space of possibilities for adaptation. Such an approach would restrict the system's inherent capacity for adaptation.

The proposed approach for integrated system design, then, has implications for existing design approaches based on CWA. Irrespective of which element or elements are of concern, it is necessary to incorporate the set of work organization possibilities in the designs of those elements. Thus, relative to existing approaches, the proposed approach would enhance the capacity of the system for adaptation by promoting structural adaptation, providing opportunities for behavioral adaptation associated with the structural possibilities, and facilitating the integration of multiple elements, such that the overall design preserves the system's underlying capacity for adaptation, across multiple actors, in a systematic fashion.

In summary, the proposed approach can be considered integrative on two levels. First, it provides a unified means for supporting adaptations in both behavior and structure. Thus, even if the focus is on an individual element, by incorporating the constraints on the possibilities for work organization in the design of that element, alongside the other constraints, the resulting design would support adaptations in both behavior and structure. Second, the approach provides a lynchpin—in the form of a common set of work organization possibilities—for integrating the design of multiple elements. This mechanism is important because simply incorporating these possibilities into the design of a single element would be conducive to supporting adaptation but insufficient. Rather, the designs of the various elements must be coordinated, across multiple actors in the system, such that the system design supports the range of possibilities for structural and behavioral adaptation in a coherent manner.

## Analysis
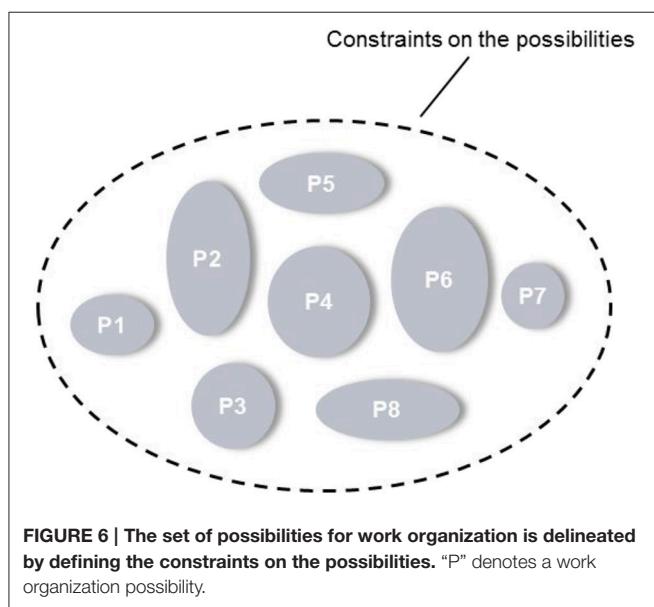
In creating an integrated system design, then, the set of work organization possibilities is a central concept in the analysis and design effort. Thus this section shows how the set of work organization possibilities may be defined, while the next section shows how these possibilities may be utilized in design.

The precise aim of the analysis phase is to demarcate the set of possibilities for work organization in a system irrespective of the situation. Thus the possibilities must be defined in a formative

manner, such that they are not limited to particular conditions but are relevant to any situation, even those that cannot be anticipated. Consequently, designs can be developed to support worker adaptation to a variety of conditions, including novel events. The key question then is how the set of all possible work structures in a system may be identified without consideration of the full set of circumstances in which they may be implemented, as all of these circumstances cannot be predicted a priori.

The essence of the approach is encapsulated in **Figure 6**. Basically, this figure shows that the set of possibilities for work organization in a system can be delineated independently of the situation by defining the constraints on the possibilities, rather than describing the possibilities themselves. As will be demonstrated in the following discussion, these constraints can be identified by analyzing the *limits* placed on the distribution of *work demands* across actors by the *criteria* that govern shifts in work organization, as these criteria will constrain the structures actors can adopt.

It is important to appreciate that the criteria that dynamically govern shifts in work organization exclude certain work structures from consideration altogether. This point is not recognized explicitly by either Rasmussen et al. (1994) or Vicente (1999). Depending on the access actors have to information or controls, for instance, only certain ways of distributing the work demands across actors will be possible in the system regardless of the situation. Likewise, based on organizational policies or the competencies of actors, only particular work arrangements will be permissible or feasible at any point in time. Thus the criteria exclude certain work structures outright, as well as constraining the structures that are suitable under particular conditions, thereby dynamically governing shifts in work organization. Consequently, by amalgamating the criteria with the work demands of the system to identify the structures that are to be excluded altogether, the set of possibilities for work organization in the system may be circumscribed.



**FIGURE 6 | The set of possibilities for work organization is delineated by defining the constraints on the possibilities.** "P" denotes a work organization possibility.

In an idealized implementation of the approach, then, the first step is to define the work demands of the system with the first three dimensions of CWA, namely work domain analysis, activity analysis, and strategies analysis, consistent with a constraint-based perspective. Accordingly, the work demands of the system will be captured in the form of an abstraction-decomposition space or abstraction hierarchy, a contextual activity template, a set of decision ladders, and a set of information flow maps (**Table 1**). As an illustration, **Figure 7** presents a modified decision ladder from a set of eight that resulted from an activity analysis of the Royal Australian Air Force's future maritime surveillance aircraft (Elix and Naikar, 2008). This model represents some of the decision-making demands associated with identifying targets, such as an enemy submarine, from the aircraft. For example, the work demands involve positioning the aircraft and manipulating its various sensors to obtain certain information about the target, such as its location and characteristics, so that the target's identity can be established, even in the face of such obstacles as the environmental conditions. The basic elements of the decision ladder template are described in detail by Rasmussen et al. (1994), Vicente (1999), and Naikar et al. (2006).

Subsequently, in the social organization and cooperation dimension, the work organization criteria are applied to the work demands to demarcate the set of possibilities for work organization in the system. As indicated above, this process involves examining the limits placed on the allocation or distribution of work demands across actors by each of the criteria, irrespective of the situation. In this paper, the same six criteria observed by Rasmussen et al. (1994) and Vicente (1999) to dynamically govern shifts in work organization are utilized. In studies of two military systems, an Airborne Early Warning and Control aircraft (Naikar et al., 2003; Naikar, 2013) and the future maritime surveillance aircraft referred to earlier, no additional criteria were identified. However, it is possible that other criteria may be relevant for different systems.

The limits on the possibilities for work organization can be identified by considering the following kinds of question in relation to the work demands captured in the various CWA models:

- *Compliance*: Does the need for compliance with policies or regulations constrain how the work demands can be allocated or distributed across actors?
- *Safety and reliability*: Does the need for safety or reliability place constraints on the allocation or distribution of work demands?
- *Access to information/controls*: Does the access actors have to information or controls constrain the allocation or distribution of work demands?
- *Coordination*: Does the need for feasible coordination requirements place constraints on how the work demands can be allocated or distributed?
- *Competencies*: Does the need for feasible competency requirements constrain the allocation or distribution of work demands?
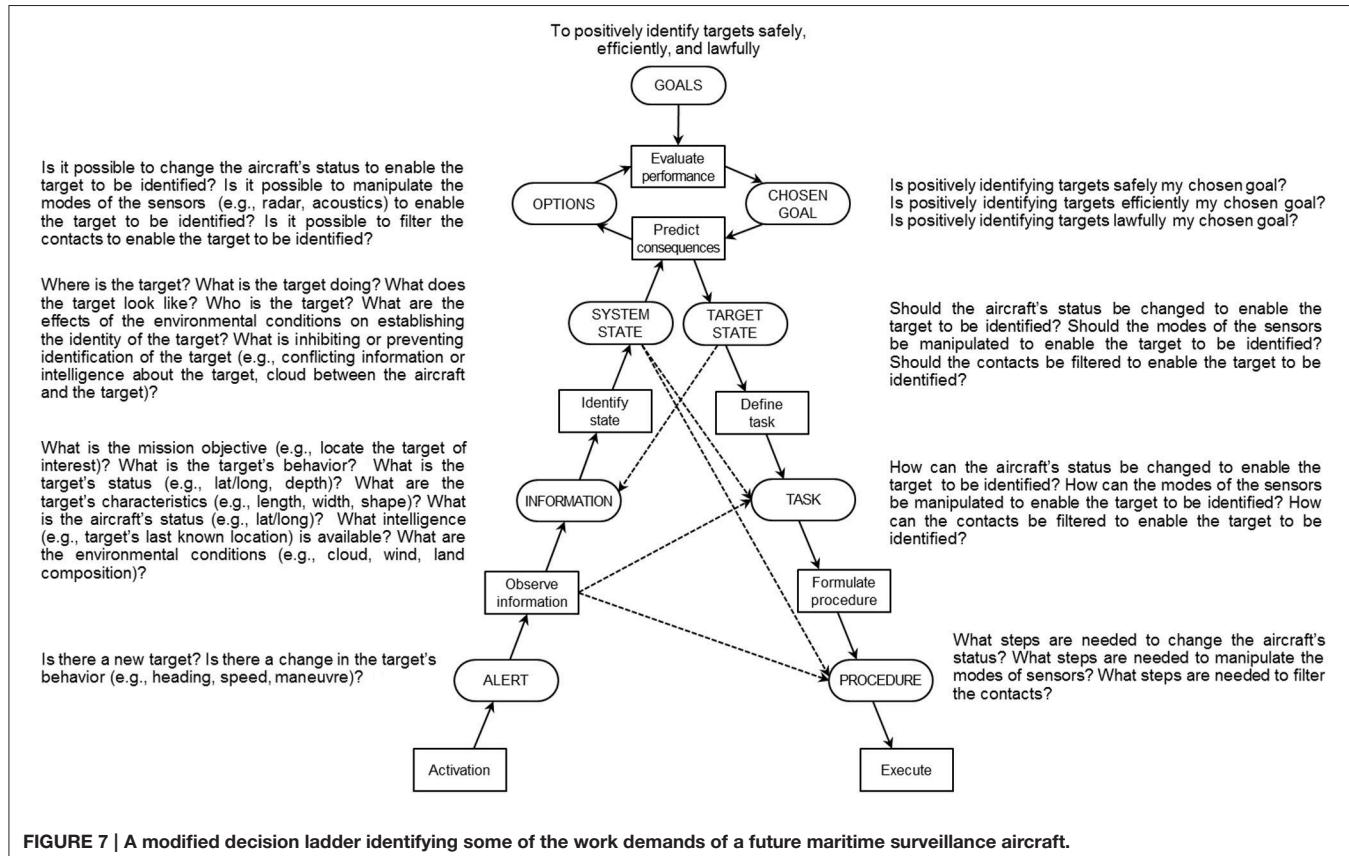
FIGURE 7 | A modified decision ladder identifying some of the work demands of a future maritime surveillance aircraft.

● *Workload*: Does the need for manageable workload constrain how the work demands can be allocated or distributed across actors?

For example, in the case of the maritime surveillance aircraft, the need for compliance with organizational regulations constrains the captaincy of the aircraft to one of the flying crew rather than tactical crew. Therefore any work demand requiring the authority of the captain, such as the arming of weapons, must be allocated to one of the flying crew (**Figure 8A**). Furthermore, the safety and reliability criterion constrains the responsibility of piloting the aircraft to two people, even though a single person would have the capacity to handle this responsibility. Consequently any work demand associated with piloting the aircraft must be allocated to at least two actors (**Figure 8B**). Third, the criterion of access to information or controls constrains the allocation of any work demand requiring a window, such as the sighting of targets, to actors in the flight deck or at observer stations in the cabin (**Figure 8C**). In addition, this criterion constrains the control of four sensor systems (i.e., the radar, electro-optical/infrared, electronic support measures, and acoustics sensors) for detecting, tracking, and identifying targets to actors at any of six workstations in the cabin (**Figure 8D**). Finally, while the criterion of minimizing coordination would constrain the operation of all of the sensors to a single actor (**Figure 8E**), the requirement for crew members to develop the necessary competencies within a reasonable timeframe and have

a manageable workload would result in the allocation of these sensors to more than one actor (**Figure 8F**).

It is important to emphasize that the criteria are applied to the work demands independently of the situation. This means that the limits that are identified on the allocation or distribution of work demands must hold regardless of the circumstances or, in other words, be relevant to any situation. From a practical perspective, then, when analysts step through the process of applying the criteria to the work demands, they are likely to find that while certain possibilities for work organization can be excluded outright on this basis, there are many remaining possibilities and which of these possibilities will be adopted by actors cannot be established independently of the situation.

In some cases, these "ambiguities" may be resolved by analysts in relation to certain classes of situation, such as the work situations in a contextual activity template (Naikar et al., 2006), which may be informative for design but limited in that there is no accounting for unanticipated events or unexpected variations in situations. However, in many cases, these uncertainties can only be resolved by actors in relation to the particularities of a situation, given that these cannot always be predicted a priori. For example, although actors may generally seek to minimize coordination requirements in enacting organizational structures to deal with events, there may be circumstances in which they adopt work structures involving greater coordination because of the workload of particular actors at that point in time. Therefore,
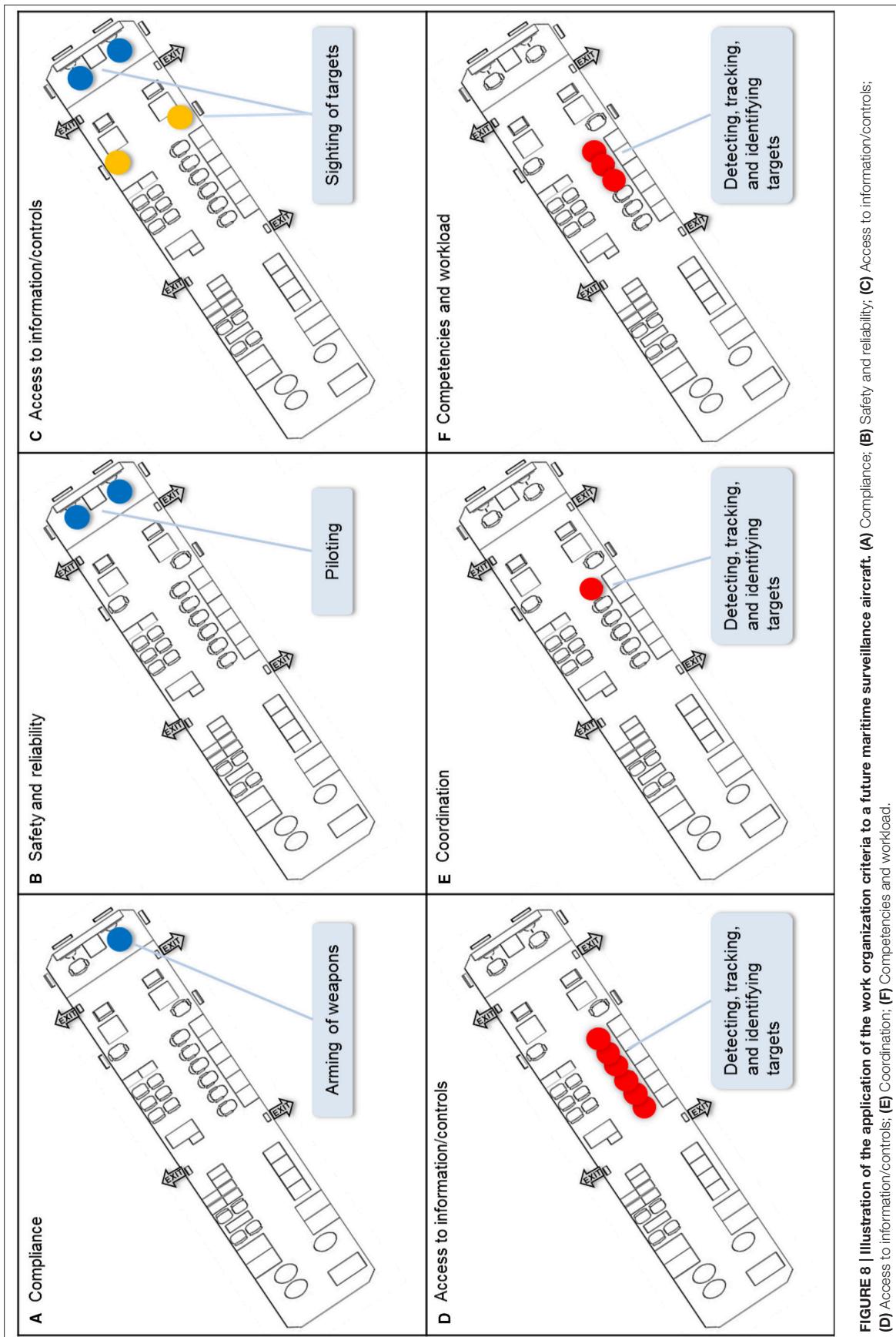
**FIGURE 8 | Illustration of the application of the work organization criteria to a future maritime surveillance aircraft. (A)** Compliance; **(B)** Safety and reliability; **(C)** Access to information/controls; **(D)** Access to information/controls; **(E)** Coordination; **(F)** Competencies and workload.

often the criterion of coordination will not result in limits on work organization being established conclusively. The same applies to the workload criterion in that there may be times when actors adopt organizational structures involving a high workload for some actors, although they may generally seek a manageable workload for all actors.

Hence, in applying the criteria to the work demands, it is important to focus on those limits that cannot be broken, irrespective of the situation. This means that the boundaries on work organization will stem largely from the criteria of compliance, safety and reliability, access to information and controls, and competencies, as event-independent limits may be derived more readily from these criteria. For instance, the access actors have to some kinds of information or controls will not vary according to situation. Similarly, many organizational regulations will hold across all situations. Nevertheless, despite these constraints, actors will still have many degrees of freedom for action, such that any of the criteria may be invoked online and in real time by actors to enact organizational structures that are suitable given the circumstances. Thus the criteria will still govern shifts in work organization dynamically.
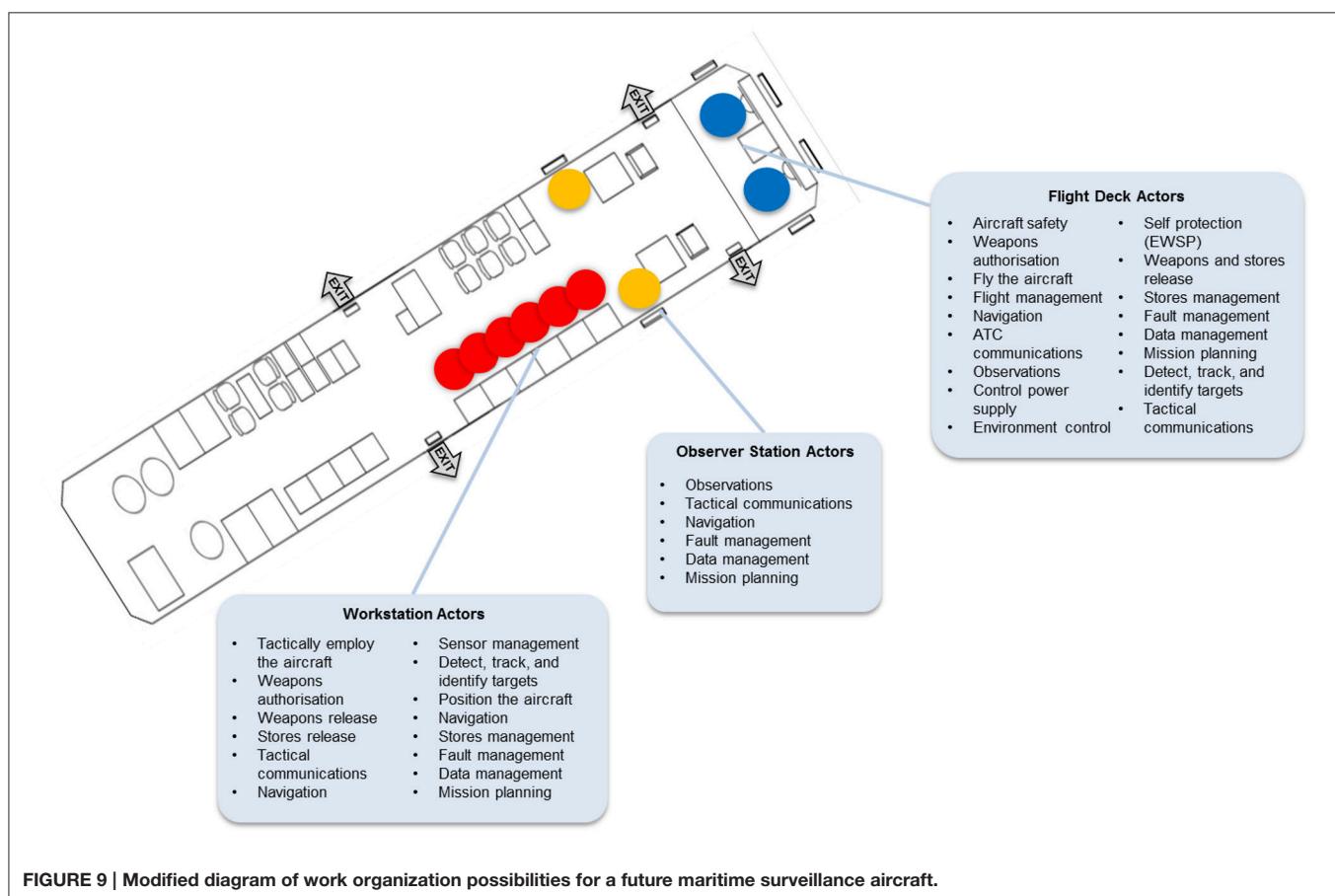
Once the criteria have been applied to the work demands to identify the limits on their distribution, it is possible to create a diagram of work organization possibilities for the actors in the system. **Figure 9** shows a modified representation of the
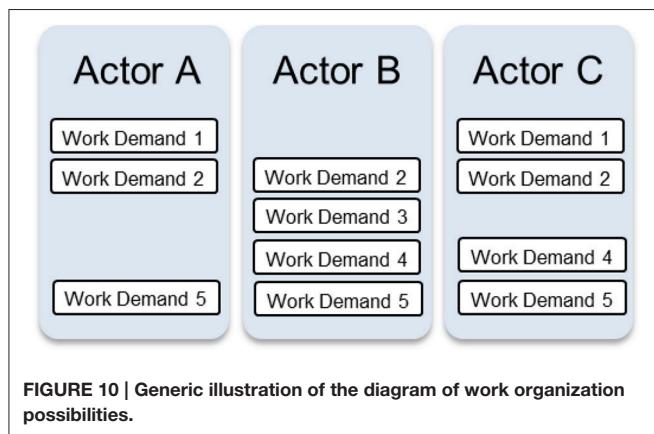
resulting diagram for the future maritime surveillance aircraft (The full diagram cannot be reproduced here because of space limitations and proprietary restrictions). This figure identifies some of the actors in the system, in terms of their positioning at particular stations on the aircraft, and provides an event-independent representation of the work demands for which these actors can be responsible.

For the sake of simplicity, **Figure 10** depicts the diagram of work organization possibilities in a generic form. In the following discussion, this figure will be drawn on to highlight some key features of this formative representation. Some examples from the maritime surveillance aircraft will also be provided.

## Can Be, Not Will Be

An important feature of the WOP diagram is that it results in an understanding of the set of work demands for which an actor *can* be responsible. Which work demands an actor *will* be responsible for at any point in time is situation-dependent, such that the responsibilities of actors could vary over time. For example, initially Actor A could be responsible for Work Demand 2 but subsequently this responsibility could be assumed by Actors B or C (**Figure 10**). In the same way, initially Actor B could be responsible for Work Demands 2, 3, 4, and 5 and subsequently for just Work Demand 3.



**FIGURE 9 | Modified diagram of work organization possibilities for a future maritime surveillance aircraft.**

FIGURE 10 | Generic illustration of the diagram of work organization possibilities.

In the case of the maritime surveillance aircraft, both of the flying crew can take responsibility for the work demands associated with navigating the aircraft (**Figure 9**). Therefore, the responsibility for these work demands might shift between these actors, depending on the situation, such that at one point in time one of these actors has this responsibility, whereas at another point in time the other actor has this responsibility. Moreover, actors at the observer stations and workstations in the aircraft's cabin can also contribute to some of the navigation work demands, such that the responsibilities for these activities could shift to these actors on certain occasions. In the same way, the actors on the flight deck have access to certain information obtained by the aircraft's sensor systems, so that, when necessary, they can contribute to some of the work demands associated with detecting, tracking, and identifying targets, either alongside or instead of the actors at the six workstations. Finally, each of the actors at the six workstations has access to the information and controls necessary for commanding the mission, which means that the responsibilities for the associated work demands can shift between these actors if required.

## Constraints vs. Possibilities
Another feature of the WOP diagram is that it demarcates the set of possibilities for work organization in a system, or the constraints on the possibilities, but it does not portray each possibility. In other words, it depicts the fundamental boundaries on the allocation or distribution of work demands, from which the various possibilities may be derived, but it does not elucidate each possibility. This distinction may be clarified further with a simple example. **Figure 10** shows that Actors A and C can take responsibility for Work Demand 1. These are the constraints or boundaries on the possibilities. Given these constraints, the possibilities are: Actor A has this responsibility, Actor C has this responsibility, or Actors A and C share this responsibility. Thus, in a given situation, if the safety criterion is emphasized, for instance, one of these possibilities may be adopted, whereas if priority is given to the criterion of workload sharing, another possibility may be adopted. Which possibility is adopted will depend on the details of the situation, which may not always be known a priori, such that the problem can only be resolved online and in real time by actors.

In the case of the maritime surveillance aircraft, the responsibility for the sighting of targets can be assumed only by actors positioned at a window and thus at four stations on the aircraft—two flight deck stations and two observer stations (**Figure 9**). These are the constraints on the allocation of this work demand. However, within these constraints, there are numerous possibilities for work organization. If one considers just the two flying crew, the possibilities are that one of the flying crew has this responsibility, the other flying crew has this responsibility, or both flying crew share this responsibility. If one includes the actors at the two observer stations, one at each station, the number of possibilities increases to 15. Moreover, if one considers the fact that each of the four stations could accommodate more than one actor, if necessary, the possibilities are considerably greater. As an example, if there is an electrical failure, such that none of the sensors can be used for detecting targets, more than one actor might be positioned at each of the four stations to increase the chances of finding the target. The WOP diagram accounts for these possibilities but it does not describe each possibility.

## Computable, But Unnecessary
Clearly, then, depending on the scale of the system and the level of granularity at which the work demands are modeled, the number of possibilities may be very large. In the case of the future maritime surveillance aircraft, for example, a rough counting revealed the number of possibilities to be in the order of $10^{27}$. However, while it may not be impossible to compute all of the possibilities, it is unnecessary to do so. That is, to support adaptation, designs must simply take into account the constraints on the possibilities. As long as a design considers the set of work demands for which actors *can* be responsible, actors *will* be able to handle those work demands effectively if and when the need arises. For instance, the interface designs at the various stations on the maritime surveillance aircraft need only accommodate the set of work demands for which actors positioned at those stations can be responsible, as represented in the WOP diagram (**Figure 9**). As a result, actors will be able to implement any one of the possibilities out of the full set if necessary.

## Emergent, Not Planned a Priori
Lastly, despite the fact that the work organization possibilities may be computed or described, the possibilities are regarded as emergent, consistent with the observations of Rochlin et al. (1987). First, the number of possibilities for a complex system is likely to be so large that it is not feasible for all of the possibilities to be considered meaningfully by analysts or designers. Certainly, this was found to be the case with the future maritime surveillance aircraft. Therefore, the possibilities for work organization can only be enacted meaningfully *in situ* by actors responding to local contingencies. Furthermore, although the work organization possibilities can be computed at some level, all of the details of these possibilities, including the local interactions between actors in the system, cannot be known or pre-specified. In fact, each fundamental possibility may have many new properties as it is enacted *in situ* by actors each time. Finally, the possibilities are regarded as emergent because

it cannot be planned a priori which of the possibilities will be appropriate in unanticipated situations, as the details of these events cannot be known ahead of time. Even in situations that are regarded as familiar, there are likely to be many small variations in context that make prediction difficult. Therefore, typically only actors can enact sensibly particular possibilities for work organization from the fundamental set, in response to the local context, and thus finish the design.
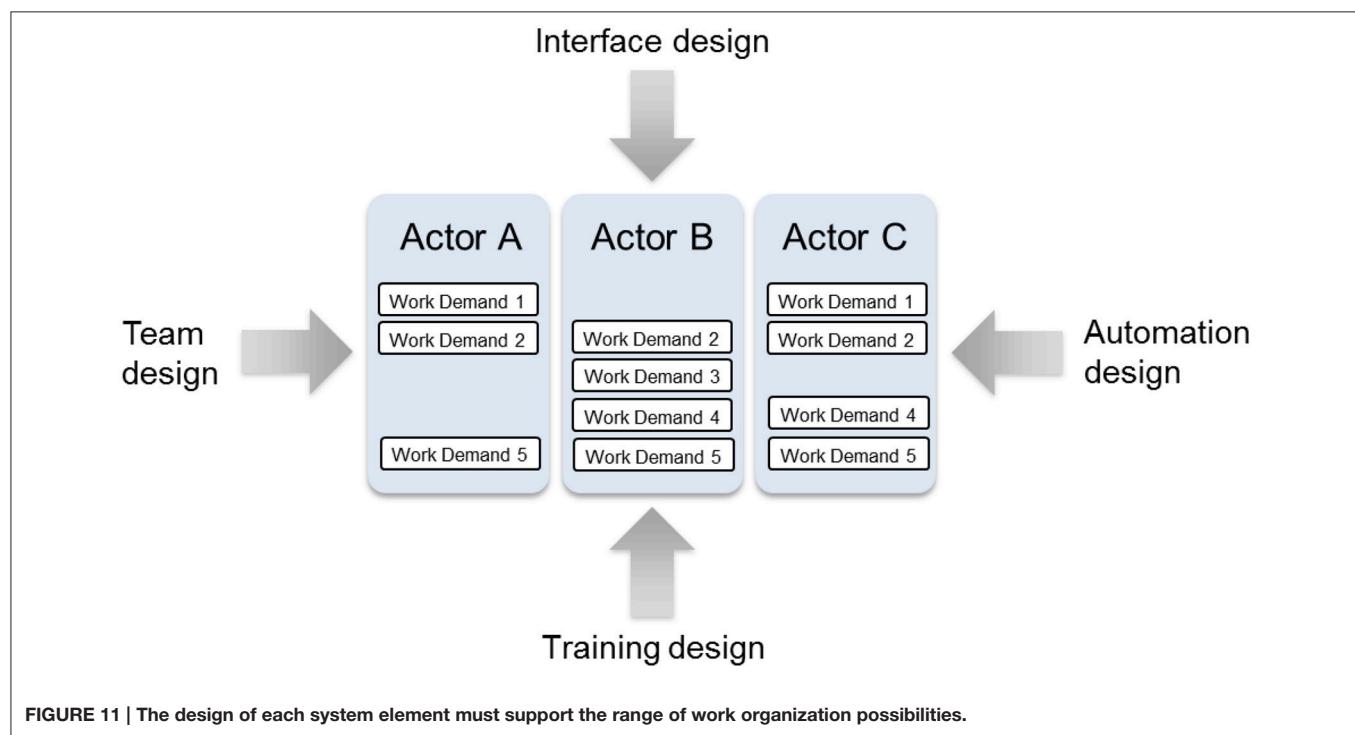
## Design

Following the analytic effort to create a diagram of work organization possibilities for the actors in the system is the design phase. This section discusses how this diagram, or the set of work organization possibilities, can be utilized in design. First, the overarching design objectives are described. Subsequently, the design of particular elements is considered, specifically by illustrating how existing design approaches for individual elements, with complementary objectives, may be extended for the purposes of creating an integrated system design.

In the proposed approach, the aim of design—of each of the system elements—is to support the set of work organization possibilities, as identified in the WOP diagram. This idea is encapsulated in **Figure 11**. This figure conveys that the team and interface designs should be such that the range of possibilities for work organization can be adopted. Similarly, the automation and training designs should support this set of possibilities. In this way, the proposed approach anchors the design of these and other elements to the organizational constraints, so that multiple actors are supported in adapting their structure as well as their behavior in a coordinated manner, regardless of the situation.

Key to this principal objective is the idea that the design of each element should not artificially constrain the capacity of the system for adaptation. That is, the designs should not incorporate extraneous constraints, or constraints beyond those fundamental to the system, such that they limit unnecessarily the possibilities for work organization. For example, the roles of actors in a team should not be so construed that the team design eliminates reasonable alternatives for distributing the work demands across actors. Likewise, the information content of the displays made available to actors should not constrain the responsibilities each can adopt, by presenting information limited to a relatively narrow range of work demands, such that the set of possibilities for work organization in the system is constricted without reason.

Also central to this design perspective is the idea that the design of each element should seek to promote the capacity of the system for adaptation by alleviating any challenges or difficulties associated with realizing or executing the possibilities. As an example, if the suite of work demands for which an actor can be responsible requires considerable competencies, consideration should be given to how the learning demands can be managed through design, perhaps of the team and training program. It may be feasible, for instance, for the actor to serve as a deputy to a more senior position within the team, following some basic instruction, such that the full set of competencies for the job can be matured gradually through on-the-job training. Alternatively, if the combination of work demands for which an actor can be responsible entails substantial workload, emphasis could be placed on reducing the cognitive effort required for particular activities through the design of the display or automated decision aids. Finally, if the array of work demands for which an actor can be responsible involves significant coordination with other



**FIGURE 11 | The design of each system element must support the range of work organization possibilities.**

actors, consideration could be given to how the communication demands can be eased through the design of the workspace layout or collaboration technologies.

Within this overarching framework, complementary design approaches may be extended to develop the various elements. For example, ecological interface design (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1990, 1992; Burns and Hajdukiewicz, 2004; Bennett and Flach, 2011) can be extended to support the development of the interface by including the delineation of work organization possibilities. As this approach stands currently, a work domain analysis [and activity analysis, if one assumes the process Bennett and Flach (2011) describe] is conducted with the goal of identifying information requirements for displays for an actor, or actors, in the system. With the view of creating an integrated system design, however, a work domain analysis would be conducted also with the intention of demarcating the set of work organization possibilities. The work domain model would still be used to derive information requirements for displays, much like in the original approach, but with a key difference being that the information requirements would be based on the set of work demands for which each actor can be responsible, as established in the WOP diagram. These information requirements would be incorporated into the displays in a way that supports skill-based, rule-based, and knowledge-based behavior, as consistent with the original approach. The resulting interface, then, would provide workers with the information necessary for fulfilling the range of responsibilities they can adopt, not just those they are allocated or usually adopt, for instance in recurring classes of situation.

Notably, both Rasmussen et al. (1994) and Vicente (1999) recognize the importance of the work organization for design. This is reflected prominently in the fact that CWA, as described in these texts, includes a social organization and cooperation dimension. Moreover, Vicente, as a case in point, states explicitly that "the division and coordination of work determines what information content actors need to perform their duties" and that "making decisions about how work demands should be divided up has important implications for the identification of relevant information content" (p. 254). However, the existing approach for ecological interface design does not address how the implications of the work organization for the interface design can be derived systematically and, more specifically, in a formative manner. Therefore, while this approach can support adaptations in actors' behavior, in its current form it does not necessarily accommodate adaptations in their structure nor support the corresponding behavioral possibilities. Furthermore, it does not necessarily facilitate the integration of the interface with other elements, across multiple actors in the system, such that the range of possibilities for adaptation is supported in a coherent manner by the system design. The approach proposed here provides a means for addressing the organizational constraints in the design of the interface element.

Similarly, an existing approach utilizing CWA for team design (Naikar et al., 2003; Naikar, 2013) can be expanded to incorporate the delineation of work organization possibilities. While this approach does attempt to accommodate flexibility in the work structure through the team design, it is limited in its capacity

to promote adaptation. Specifically, for a given system, work domain analysis and activity analysis are used to explore the feasibility of alternative team concepts, or alternative possibilities for work organization, in the context of plausible scenarios. On this basis, the strengths and limitations of the alternative concepts are identified, and requirements are generated for a new team design with the intent of capitalizing on the various possibilities. One limitation of this approach is that it relies on pre-conceived team concepts that are not necessarily constraint-based, such that it limits artificially the work organization possibilities that are considered. Moreover, the alternative team concepts are considered initially in the context of plausible situations, albeit both common and exceptional ones. Notably, this approach does involve generalizing beyond the particular situations examined, by translating the work demands in the scenarios into recurring work functions from the contextual activity template and by examining the impact of the alternative team concepts on the work domain constraints, which are relevant to a broad range of events including unforeseen ones. Nevertheless, a more parsimonious solution would be beneficial.

The approach proposed here provides a way of using CWA to generate the set of possibilities for work organization independently of the situation, within the constraints of the system, as captured in the WOP diagram. As a result, the requirements for the team design, such as the number, roles, and hierarchical levels of people in the team, can be defined in light of the suite of work demands actors can fulfill, regardless of the circumstances they find themselves in. Aside from accommodating greater possibilities for behavioral and structural adaptation, this team design can be integrated with other elements, across multiple actors, such that the possibilities are supported uniformly throughout the system.

An approach for using CWA for training design (Naikar and Sanderson, 1999; Naikar, 2013) also can be broadened to take into account the set of work organization possibilities. The current approach seeks to foster adaptation by promoting the design of training systems that offer the same possibilities for action that are afforded by the work environment or work domain. For example, a simulator with parallel means-ends structure to the work domain, or structural means-ends fidelity, will allow workers to exploit the same means-ends relations that are available in their actual work context. Hence, with the aid of a suitable training program, workers can become proficient in exploiting flexibly multiple system means, or resources, to fulfill multiple system ends, or purposes, such that they can respond in novel or adaptive ways to handle abnormal or unpredictable situations. Thus, in this approach, work domain analysis is central for defining the features or characteristics of training equipment or devices, such as high-fidelity simulators, whereas the remaining CWA dimensions provide a strong foundation for defining complementary training programs, although each dimension can inform either problem (see also Lintern and Naikar, 2000; Jenkins et al., 2011).

These ideas may be expanded for the purposes of creating an integrated system design. In developing training equipment or devices, consideration must be given to those physical and intentional features of the work domain that constrain or afford

the work organization possibilities that are available to workers in their actual work context. In the case of high-fidelity simulators particularly, it may be desirable to recreate these properties so that workers have the same possibilities for work organization during training that are available to them otherwise. Similarly, training programs should give consideration to the full set of work demands that actors can assume responsibility for in their actual work context, as specified in the WOP diagram, so that workers are more suitably prepared for exploiting the range of possibilities for adaptation.

Finally, frameworks for automation design that are concerned principally with human-automation coordination (Dekker and Woods, 2002; Klein et al., 2004; Hollnagel and Woods, 2005; Woods and Hollnagel, 2006; Bradshaw et al., 2013) can be extended to take into account the set of work organization possibilities. Although these frameworks do not intrinsically involve the use of CWA, they are consistent with a constraint-based perspective in some respects. These frameworks recognize that the conventional preoccupation with the allocation of functions between humans and machines is limited and that the primary question of concern is not what level of autonomy or control is to be assigned to the human vs. the machine but rather, given the capabilities of the automation, how to support the interaction that would necessarily be required between humans and machines, if the capabilities of the automation are to be exploited.

This viewpoint aligns with the proposed approach for integrated system design. In relation to automation design specifically, the proposed approach recognizes that rather than focusing on pre-specifying a limited number of schemes for allocating work demands between humans and machines, which would inevitably be limited to anticipated events, it is necessary to identify the work demands that can be handled by the automation, alongside the human actors, irrespective of the situation. Subsequently, the interaction demands associated with the set of possibilities for work organization, encompassing both humans and machines, can be supported through design. Therefore, in the analysis phase, the automated agents can be treated as actors, as originally recognized by Rasmussen et al. (1994) and Vicente (1999), such that the set of work organization possibilities encompasses the potential distributions of work demands across human and machine actors. As per the earlier discussion, which possibility is adopted by humans, as only humans can take ownership of problems (Bradshaw et al., 2013), is dependent on the situation, such that sometimes the work structure might include both human and machine actors and sometimes not. Hence, the key implication for the design phase is the need to ensure that any one of the possibilities encapsulated in the WOP diagram can be implemented effectively, specifically by supporting the interaction demands associated with the range of work arrangements.

It is important to point out that the preceding discussion does not address all of the nuances in the implications of the integrated system design approach for the design of individual components. Nor does it address the full range of elements. This is beyond the scope of this paper. Rather, the intent has been

to provide an illustration of how some existing, complementary design approaches for individual elements can be extended for the purposes of creating an integrated system design. Through these extensions the design of multiple elements can be anchored to, and thus coordinated around, the set of possibilities for work organization, such that the system design supports the structural and behavioral opportunities for adaptation systematically across multiple actors.

## CONCLUSION

To conclude, this paper has proposed an approach for integrated system design, based on extensions of CWA. This approach recognizes that to promote the capacity of sociotechnical systems for adaptation, the designs of the various elements must be integrated, such that workers are supported in adapting their structure as well as their behavior in a coherent manner. To this end, the approach proposes the set of possibilities for work organization in a system as the central mechanism for coordinating the design of multiple elements across multiple actors. Accordingly, the paper demonstrates how the set of work organization possibilities may be demarcated independently of the situation and how the resulting diagram of work organization possibilities may be utilized in design. Relative to existing analysis and design frameworks, this approach has the potential to enhance a system's capacity for adaptation by accommodating possibilities for structural adaptation across a variety of situations including unforeseen ones, supporting opportunities for behavioral adaptation associated with those structural possibilities, and facilitating the integration of multiple elements such that the system design supports the range of possibilities for adaptation, across multiple actors, in a systematic fashion.

As noted at the outset, the rationale for the proposed approach rests on the assumption that the principal design objective for sociotechnical systems should be that of facilitating successful adaptation, as these systems are open to changing conditions, including unanticipated events, which pose the most substantive threats to their viability. Moreover, supporting worker adaptation in everyday and novel situations is important not just for preserving system safety but also for promoting organizational productivity and workers' health. As ongoing adaptation to dynamic and unforeseen conditions is a highly exacting mode of operating, workers should be supported—deliberately and systematically—in adapting to the demands of the entire range of events through the system design. In particular, as consistent with considerable empirical evidence, the system design should support workers in adapting both their organization and behavior to the changing circumstances.

The proposed approach recognizes that current approaches for work analysis and design are limited in their capacity to support adaptation, for the most part, because they are focused on specifying optimal ways of performing work or describing existing ways of performing work under particular conditions, whether these are familiar, recurring, or anticipated.

The CWA framework circumvents this limitation by focusing on the constraints on actors, rather than on the details of their work practices, as these constraints are relevant to both known and novel events. However, while this framework can support adaptation, as has been demonstrated empirically, the adaptations are constrained primarily to actors' behavior and do not necessarily extend to their work structure and the corresponding behavioral possibilities. A related problem is that CWA is limited in its ability to support integrated system design. Given that the design of various elements must be integrated across multiple actors in the system, understanding the possibilities for structural adaptation is necessary, and CWA, in its present form, does not provide a means by which this can be achieved comprehensively. The approach for integrated system design addresses these issues in the manner summarized above.

In closing, it is important to acknowledge that while the approach for integrated system design described in this paper has been demonstrated conceptually, its viability has not been fully established. Considerably more work is necessary to achieve this goal (Naikar and Elix, 2016b). A key objective of future research should be to validate the various ideas constituting this approach, either through experimental studies or case studies. Another important question to be addressed relates to the feasibility of implementing the complete approach, or aspects of it, in industrial contexts. By providing a comprehensive description of the approach for integrated system design, this paper enables these critical objectives to be pursued.

# AUTHOR CONTRIBUTIONS

# ACKNOWLEDGMENTS

# REFERENCES

32nd Army Air Missile Defense Command (2003). *Patriot Missile Defense Operations during Iraqi Freedom.* Washington DC: U.S. Army.

Ashby, W. R. (1956). *An Introduction to Cybernetics.* London: Chapman & Hall.

Bennett, K. B., and Flach, J. M. (2011). *Display and Interface Design: Subtle Science, Exact Art.* Boca Raton, FL: CRC Press.

Bennett, K. B., Posey, S. M., and Shattuck, L. G. (2008). Ecological interface design for military command and control. *J. Cogn. Eng. Decis. Making* 2, 349–385. doi: 10.1518/155534308X377829

Bigley, G. A., and Roberts, K. H. (2001). The incident command system: high-reliability organizing for complex and volatile task environments. *Acad. Manag. J.* 44, 1281–1299. doi: 10.2307/3069401

Bogdanovic, J., Perry, J., Guggenheim, M., and Manser, T. (2015). Adaptive coordination in surgical teams: an interview study. *BMC Health Serv. Res.* 15:128. doi: 10.1186/s12913-015-0792-5

Borst, C., Suijkerbuijk, H. C. H., Mulder, M., and van Paassen, M. M. (2006). Ecological interface design for terrain awareness. *Int. J. Aviat. Psychol.* 16, 375–400. doi: 10.1207/s15327108ijap1604_3

Bradshaw, J. M., Hoffman, R. R., Johnson, M., and Woods, D. D. (2013). The seven deadly myths of "autonomous systems". *IEEE Intell. Syst.* 28, 54–61. doi: 10.1109/MIS.2013.70

Bureau of Enquiry and Analysis for civil aviation safety (BEA) (2002). *AF4590 Final Report, Rep. Bureau Enquêtes Et Analyses,* 25, 2000.

Bureau of Enquiry and Analysis for civil aviation safety (BEA) (2012). *AF447 Final Report, Rep. Bureau Enquêtes Et Analyses,* June 1, 2009.

Burns, C. M., and Hajdukiewicz, J. R. (2004). *Ecological Interface Design.* Boca Raton, FL: CRC Press.

Burns, C. M., Kuo, J., and Ng, S. (2003). Ecological interface design: a new approach for visualizing network management. *Comp. Netw.* 43, 369–388. doi: 10.1016/S1389-1286(03)00287-1

Christoffersen, K., Hunter, C. N., and Vicente, K. J. (1996). A longitudinal study of the effects of ecological interface design on skill acquisition. *Hum. Factors* 38, 523–541. doi: 10.1518/001872096778701917

Czaja, S. J. (1997). "Systems design and evaluation," in *Handbook of Human Factors and Ergonomics*, ed G. Salvendy (Hoboken, NJ: John Wiley & Sons), 17–40.

Dekker, S. (2003). Failure to adapt or adaptations that fail: contrasting models on procedures and safety. *Appl. Ergon.* 34, 233–238. doi: 10.1016/S0003-6870(03)00031-0

Dekker, S. W. A., and Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation co-ordination. *Cogn. Technol. Work* 4, 240–244. doi: 10.1007/s101110200022

Dörner, D. (1996). *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations.* New York, NY: Basic Books.

Eason, K. (2014). Afterword: the past, present and future of sociotechnical systems theory. *Appl. Ergon.* 45, 213–220. doi: 10.1016/j.apergo.2013.09.017

Elix, B., and Naikar, N. (2008). "Designing safe and effective future systems: a new approach for modelling decisions in future systems with cognitive work analysis," in *Proceedings of the 8th International Symposium of the Australian Aviation Psychology Association* (Sydney, NSW: Australian Aviation Psychology Association).

Emery, F. E., and Trist, E. L. (1965). The causal texture of organizational environments. *Hum. Relations* 18, 12–32. doi: 10.1177/001872676501800103

Gerson, E. M., and Star, S. L. (1986). Analyzing due process in the workplace. *ACM Trans. Off. Inf. Syst.* 4, 257–270. doi: 10.1145/214427.214431

Goteman, Ö., and Dekker, S. (2001). "Flight deck call-outs and automation awareness," in *Proceedings of the Conference on Human Factors and Safety in Aviation* (Lund: Swedish Centre for Aviation Research and Development), 64–72.

Ham, D. H., and Yoon, W. C. (2001). The effects of presenting functionally abstracted information in fault diagnosis tasks. *Reliability Eng. Syst. Safety* 73, 103–119. doi: 10.1016/S0951-8320(01)00053-9

Herzog, S. (2011). Revisiting the Estonian cyber attacks: digital threats and multinational responses. *J. Strateg. Security* 4, 49–60. doi: 10.5038/1944-0472.4.2.3

Hoffman, R. R., and Woods, D. D. (2011). Beyond Simon's slice: five fundamental trades-offs that bound the performance of macrocognitive work systems. *IEEE Intell. Syst.* 26, 67–71. doi: 10.1109/MIS.2011.97

Hollnagel, E., Pariès, J., Woods, D. D., and Wreathall, J. (2011). *Resilience Engineering in Practice: A Guidebook.* Farnham, UK: Ashgate.

Hollnagel, E., and Woods, D. D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering.* Boca Raton, FL: CRC Press.

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

Hoppe, J., and Popham, P. (2007). Complete failure of spinal anaesthesia in obstetrics. *Int. J. Obstet. Anesth.* 16, 250–255. doi: 10.1016/j.ijoa.2006.12.005

Hori, S., Vicente, K. J., Shimizu, Y., and Takami, I. (2001). "Putting cognitive work analysis to work in industry practice: integration with ISO13407 on human-centered design," in *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (Santa Monica, CA: HFES), 429–433.

Institute for Safe Medication Practices (2007). *Fluorouracil incident Root Cause Analysis.* Toronto, ON: ISMP. Available online at: http://www.ismp-canada.org/download/reports/FluorouracilIncidentMay2007.pdf

Jamieson, G. (2007). Ecological interface design for petrochemical process control: an empirical assessment. *IEEE Trans. Syst. Man Cybernet. Part A Syst. Hum.* 37, 906–920. doi: 10.1109/TSMCA.2007.897583

Jenkins, D. P., Stanton, N. A., Salmon, P. M., and Walker, G. H. (2011). A formative approach to developing synthetic environment fidelity requirements for decision-making training. *Appl. Ergon.* 42, 757–769. doi: 10.1016/j.apergo.2010.12.003

Karasek, R., and Theorell, T. (1990). *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life.* New York, NY: Basic Books.

Kirwan, B., and Ainsworth, L. K. (eds.). (1992). *A Guide to Task Analysis: The Task Analysis Working Group.* Boca Raton, FL: CRC Press.

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

Lau, N., Jamieson, G. A., Skraaning, G. Jr., and Burns, C. M. (2008). Ecological interface design in the nuclear domain: an empirical evaluation of ecological displays for the secondary subsystems of a boiling water reactor plant simulator. *IEEE Trans. Nucl. Sci.* 55, 3597–3610. doi: 10.1109/TNS.2008.2005725

Leveson, N. G. (1995). *Safeware: System Safety and Computers.* Reading, MA: Addison-Wesley.

Leveson, N. G., and Turner, C. S. (1993). An investigation of the Therac-25 accidents. *Computer* 26, 18–41. doi: 10.1109/MC.1993.274940

Lintern, G., and Naikar, N. (2000). "The use of work domain analysis for the design of training systems," in *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society* (Santa Monica, CA: HFES), 198–201.

Mannan, M. S., West, H. H., and Berwanger, P. C. (2007). Lessons learned from recent incidents: facility siting, atmospheric venting, and operator information systems. *J. Loss Prev. Process Industries* 20, 644–650. doi: 10.1016/j.jlp.2007.04.015

Militello, L. G., Sushereba, C. E., Branlat, M., Bean, R., and Finomore, V. (2015). Designing for military pararescue: naturalistic decision-making perspective, methods, and frameworks. *J. Occup. Organ. Psychol.* 88, 251–272. doi: 10.1111/joop.12114

Naikar, N. (2006). "An examination of the key concepts of the five phases of cognitive work analysis with examples from a familiar system," in *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society* (Santa Monica, CA: HFES), 447–451.

Naikar, N. (2012). *Cognitive Work Analysis: Foundations, Extensions, and Challenges.* DSTO General Discussion (DSTO-GD-0680). Fishermans Bend, VIC: Air Operations Division.

Naikar, N. (2013). *Work Domain Analysis: Concepts, Guidelines, and Cases.* Boca Raton, FL: CRC Press.

Naikar, N., and Elix, B. (2016a). "A consideration of design approaches based on cognitive work analysis: system design and integrated system design," in *Proceedings of the 2016 European Conference on Cognitive Ergonomics* (New York, NY: Association of Computing Machinery).

Naikar, N., and Elix, B. (2016b). Reflections on cognitive work analysis and its capacity to support designing for adaptation. *J. Cogn. Eng. Decis. Making.* doi: 10.1177/1555343416654846. [Epub ahead of print].

Naikar, N., and Elix, B. (2015). "Designing for adaptation with cognitive work analysis: initial conceptualisation of an approach for integrated sociotechnical system design," in *Proceedings of the 19th Triennial Congress of the International Ergonomics Association.* Available online at: http://ergonomics.uq.edu.au/iea/proceedings/Index_files/papers/468.pdf

Naikar, N., Moylan, A., and Pearce, B. (2006). Analysing activity in complex systems with cognitive work analysis: concepts, guidelines and case study for control task analysis. *Theor. Issues Ergon. Sci.* 7, 371–394. doi: 10.1080/14639220500098821

Naikar, N., Pearce, B., Drumm, D., and Sanderson, P. M. (2003). Designing teams for first-of-a-kind, complex systems using the initial phases of cognitive work analysis: case study. *Hum. Factors* 45, 202–217. doi: 10.1518/hfes.45.2.202.27236

Naikar, N., and Sanderson, P. (1999). Work domain analysis for training-system definition and acquisition. *Int. J. Aviat. Psychol.* 9, 271–290. doi: 10.1207/s15327108ijap0903_5

Naikar, N., and Sanderson, P. M. (2001). Evaluating design proposals for complex systems with work domain analysis. *Hum. Factors* 43, 529–542. doi: 10.1518/001872001775870322

Park, Y., Hong, P., and Roh, J. J. (2013). Supply chain lessons from the catastrophic natural disaster in Japan. *Bus. Horiz.* 56, 75–85. doi: 10.1016/j.bushor.2012.09.008

Pawlak, W. S., and Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. *Int. J. Hum. Comput. Stud.* 44, 653–688. doi: 10.1006/ijhc.1996.0028

Perrow, C. (1984). *Normal Accidents: Living with High Risk Technologies.* New York, NY: Basic Books.

Rankin, A., Lundberg, J., Woltjer, R., Rollenhagen, C., and Hollnagel, E. (2014). Resilience in everyday operations: a framework for analyzing adaptations in high-risk work. *J. Cogn. Eng. Decis. Making* 8, 78–97. doi: 10.1177/1555343413498753

Rasmussen, J. (1968a). *Characteristics of Operator, Automatic Equipment and Designer in Plant Automation (Risø-M-808).* Roskilde: Risø.

Rasmussen, J. (1968b). *On the Reliability of Process Plants and Instrumentation Systems (Risø-M-706).* Roskilde: Danish Atomic Energy Commission, Research Establishment Risø.

Rasmussen, J. (1969). *Man-Machine Communication in the Light of Accident Records* (Report No. S-1-69). Roskilde: Danish Atomic Energy Commission, Research Establishment Risø.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering.* New York, NY: North-Holland.

Rasmussen, J. (1997). "Merging paradigms: decision making, management, and cognitive control," in *Decision Making Under Stress: Emerging Themes and Applications* eds R. Flin, E. Salas, M. Strub, and L. Martin (Aldershot, UK: Ashgate), 67–81.

Rasmussen, J., and Goodstein, L. P. (1987). Decision support in supervisory control of high-risk industrial systems. *Automatica* 23, 663–671. doi: 10.1016/0005-1098(87)90064-1

Rasmussen, J., Pejtersen, A. M., and Goodstein, L. P. (1994). *Cognitive Systems Engineering.* New York, NY: John Wiley & Sons.

Rasmussen, J., and Vicente, K. J. (1989). Coping with human errors through system design: implications for ecological interface design. *Int. J. Man Mach. Stud.* 31, 517–534. doi: 10.1016/0020-7373(89)90014-X

Read, G. J. M., Salmon, P. M., Lenné, M. G., and Jenkins, D. P. (2015a). Designing a ticket to ride with the cognitive work analysis design toolkit. *Ergonomics* 58, 1266–1286. doi: 10.1080/00140139.2015.1013576

Read, G. J. M., Salmon, P. M., Lenné, M. G., and Stanton, N. A. (2015b). Designing sociotechnical systems with cognitive work analysis: putting theory back into practice. *Ergonomics* 58, 822–851. doi: 10.1080/00140139.2014.980335

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

Reich, P. C., Weinstein, S., Wild, C., and Cabanlong, A. S. (2010). Cyber warfare: a review of theories, law, policies, actual incidents—and the dilemma of anonymity. *Eur. J. Law Technol.* 1, 1–58. Available online at: http://ejlt.org/article/view/40/58

Reising, D. V. C., and Sanderson, P. M. (1998). "Designing displays under ecological interface design: towards operationalizing semantic mapping," in *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (Santa Monica, CA: HFES), 372–376.

Reising, D. V. C., and Sanderson, P. M. (2000a). "Testing the impact of instrumentation location and reliability on ecological interface design: control performance," in *Proceedings of the Joint 14th Triennial Congress of the International Ergonomics Association/44th Annual Meeting of the Human Factors and Ergonomics Society (HFES/IEA 2000), Vol. 1* (Santa Monica, CA: HFES), 124–127.

Reising, D. V. C., and Sanderson, P. M. (2000b). "Testing the impact of instrumentation location and reliability on ecological interface design: fault diagnosis performance," in *Proceedings of the Joint 14th Triennial Congress of the International Ergonomics Association/44th Annual Meeting of the Human Factors and Ergonomics Society (HFES/IEA 2000), Vol. 3* (Santa Monica, CA: HFES), 591–594.

Rochlin, G. I., La Porte, T. R., and Roberts, K. H. (1987). The self-designing high-reliability organization: aircraft carrier flight operations at sea. *Naval War Coll. Rev.* 40, 76–90.

Sanderson, P., Naikar, N., Lintern, G., and Goss, S. (1999). "Use of cognitive work analysis across the system life cycle: from requirements to decommissioning," in *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society* (Santa Monica, CA: HFES), 318–322.

Schraagen, J. M., Chipman, S. F., and Shalin, V. L. (eds.). (2000). *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sharp, T. D., and Helmicki, A. J. (1998). "The application of the ecological interface design approach to neonatal intensive care medicine," in *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (Santa Monica, CA: HFES), 350–354.

Simon, H. (1969). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Suchman, L. A. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge UK: Cambridge University Press.

Vicente, K. J. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Vicente, K. J. (2002). Ecological interface design: progress and challenges. *Hum. Factors* 44, 62–78. doi: 10.1518/0018720024494829

Vicente, K. J., and Rasmussen, J. (1990). The ecology of human-machine systems II: mediating "direct perception" in complex work domains. *Ecol. Psychol.* 2, 207–249. doi: 10.1207/s15326969eco0203_2

Vicente, K. J., and Rasmussen, J. (1992). Ecological interface design: theoretical foundations. *IEEE Trans. Syst. Man Cybernet.* 22, 589–606. doi: 10.1109/21.156574

Weick, K. E. (1993). The collapse of sensemaking in organizations: the Mann Gulch disaster. *Admin. Sci. Q.* 38, 628–652. doi: 10.2307/2393339

Whitefield, A., Wilson, F., and Dowell, J. (1991). A framework for human factors evaluation. *Behav. Inf. Technol.* 10, 65–79. doi: 10.1080/01449299108924272

Woods, D. D., and Hollnagel, E. (2006). *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. Boca Raton, FL: CRC Press.

Xu, W., Dainoff, M. J., and Mark, L. S. (1999). Facilitate complex search tasks in hypertext by externalizing functional properties of a work domain. *Int. J. Hum. Comput. Interact.* 11, 201–229. doi: 10.1207/S15327590IJHC1103_2

# Mission Command in the Age of Network-Enabled Operations: Social Network Analysis of Information Sharing and Situation Awareness

Norbou Buchler [1]*, Sean M. Fitzhugh [1], Laura R. Marusich [1], Diane M. Ungvarsky [1], Christian Lebiere [2] and Cleotilde Gonzalez [2]

[1] U.S. Army Research Laboratory, Aberdeen Proving Ground, MD, USA, [2] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

A common assumption in organizations is that information sharing improves situation awareness and ultimately organizational effectiveness. The sheer volume and rapid pace of information and communications received and readily accessible through computer networks, however, can overwhelm individuals, resulting in data overload from a combination of diverse data sources, multiple data formats, and large data volumes. The current conceptual framework of network enabled operations (NEO) posits that robust networking and information sharing act as a positive feedback loop resulting in greater situation awareness and mission effectiveness in military operations (Alberts and Garstka, 2004). We test this assumption in a large-scale, 2-week military training exercise. We conducted a social network analysis of email communications among the multi-echelon Mission Command staff (one Division and two sub-ordinate Brigades) and assessed the situational awareness of every individual. Results from our exponential random graph models challenge the aforementioned assumption, as increased email output was associated with lower individual situation awareness. It emerged that higher situation awareness was associated with a lower probability of out-ties, so that broadly sending many messages decreased the likelihood of attaining situation awareness. This challenges the hypothesis that increased information sharing improves situation awareness, at least for those doing the bulk of the sharing. In addition, we observed two trends that reflect a compartmentalizing of networked information sharing as email links were more commonly formed among members of the command staff with both similar functions and levels of situation awareness, than between two individuals with dissimilar functions and levels of situation awareness; both those findings can be interpreted to reflect effects of homophily. Our results have major implications that challenge the current conceptual framework of NEO. In addition, the information sharing network was largely imbalanced and dominated by a few key individuals so that most individuals in the network have very few email connections, but a small number of individuals have very many connections. These results highlight several major growing pains for networked organizations and military organizations in particular.

Keywords: network organization, sociotechnical system, Pareto principle, communication exponential random graph model, homophily, degree distribution, training effectiveness

# INTRODUCTION

Advances in information and network technology continue to transform the way human organizations communicate and operate. This is evident as networked organizations are at the core of the political, military, economic, and social fabric of the twenty-first century (Castells, 2009). The same technological advances that have given rise to networked forms of organization also facilitate their study. For example, larger and larger volumes of data that characterize our "digital behaviors," including communication and collaboration, are increasingly collected by companies, governments, and researchers alike (Navaroli and Smyth, 2015). Using this digital behavior data, organizations can be characterized as social networks with nodes representing individuals and links representing the interactions between them. Many such networks are inherently complex in the sense that their structure is irregular, task- and context-specific, and dynamically evolving in time.

Over the past decade, the social sciences have seen rapid growth in research and understanding of the structure of real-world complex networks (Borgatti et al., 2009). However, the effects that operating within such complex networks have upon individual macro-cognitive processes is not well understood (Klein et al., 2003). Organizations can confer considerable advantages to information sharing as the number of potential collaborations may be virtually limitless, as is the availability of information. There are however, some potential downsides as well, as the resulting deluge of information (Gleick, 2012) can quickly overwhelm human cognitive capabilities. Understanding the relationship between network structure, human collaboration, and cognitive work processes within real organizations is a critical challenge. This is especially true in command and control domains, such as military operations, emergency response, managing safety critical systems, air traffic control, computer network defense service providers, and others. In all these naturalistic domains information from various sources and of varying quality must be quickly assimilated and shared among distributed team members to make critical decisions with potentially significant consequences.

A prevalent perspective within these domains is that increased networking capabilities lead to greater information sharing and availability of information which ultimately results in improved collaboration, organizational efficiency, and better situation awareness (SA). We explore this assumption, investigating macro-cognitive processes using data collected in a large-scale exercise of military network level operations. We focus on the relationship between information sharing and SA within a real-world networked organization.

## Network Enabled Operations and Information Fusion

The tenets of network-enabled operations (NEO; Alberts and Garstka, 2004) provide an influential conceptual framework for understanding how increased networking affects human collaboration and organizational performance within the military domain. This framework posits that communication and information sharing act as a positive feedback loop with increased information sharing resulting in greater situation awareness and mission effectiveness in military operations. From a policy perspective, enhancing information sharing within and across organizations has been and is a major priority for investment by the United States government including the Department of Defense (Alberts et al., 1999; Alberts and Garstka, 2004), Federal Emergency Management Agency, Department of Homeland Security (Department of Homeland Security, 2013), and the Federal Aviation Administration (2014). As information sharing is increasingly promoted within NEO, it becomes critical to explore and understand the relationships between information sharing, cognition, and situation awareness among the staff in these complex operational environments.

The positive effects of increased information sharing upon SA can be greatly diminished if individuals reach a state of information overload. A major tenet of the Office of the Secretary of Defense's "data-to-decision" (D2D) initiative (Swan and Hennig, 2012), and a primary challenge for military commanders and their staff is to shorten the cycle time and improve the processes of synthesizing data to information and into knowledge to support decision-making and action. Organizational performance and effectiveness are curtailed by failures or bottlenecks at any step in this D2D sequence. Effectively managing the entire process requires broad collaboration and flexibility in supporting multiple information and decision requirements. In networked organizations, however, the sheer volume and rapid pace of information and communications received and readily accessible from diverse sources and in multiple formats can quickly overwhelm individuals in the D2D pipeline. Well-designed automation and decision-support tools can provide some assistance in the D2D cycle; however, the volume of data flowing through large organizational networks is often beyond the ability of current software tools to capture, curate, and store (Salimi and Vita, 2006; White, 2012) or to process the data within a tolerable time frame (Snijders et al., 2012).

A critical process of the D2D pipeline is that of information fusion. Software tools and automation currently lack the capabilities to synthesize information in an adaptive, highly context-aware manner, which necessitates human involvement and considerable cognitive resources (Blasch et al., 2011). Many contextual factors affect the human ability to rapidly synthesize information into a coherent understanding of the current situation, including information volume, quality, and modality, the general level of risk and time-pressure in the environment, and factors operating at the level of the individual decision-maker, including cognitive load, fatigue, level of expertise, and personality traits such as need for closure and need for cognition. The concept of cognitive information fusion (Blasch et al., 2012) emphasizes the necessity and strength of the human element in order to achieve a high-level, contextual understanding of a given situation. Data fusion is a term typically used to describe computational frameworks for constructing a comprehensive, data aggregation system that processes information to support user decision-requirements (Klein, 2004), whereas cognitive information fusion explicitly emphasizes the need for human cognition and staff collaborations to integrate and rapidly make

sense of these data streams that are distributed across space and time. The outcome of proficient cognitive information fusion is high situation awareness, which we describe in detail below.

## Situation Awareness in NEO

Situation Awareness (SA) is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1995, p. 36). SA is a well-known concept in a variety of domains that require cognitive information fusion, including military operations (Endsley, 2000; Matthews et al., 2004), aviation (Kaber et al., 2002; Keller et al., 2004), air traffic control (Endsley and Kiris, 1995; Endsley and Smolensky, 1998; Hauss and Eyferth, 2003), transportation (Zheng et al., 2004) and many others. The three-level model of SA proposed by Endsley (1995) is perhaps the most common model of SA (other models include those discussed in Smith and Hancock, 1995 and Bedny and Meister, 1999). Endsley's model depicts SA as an essential input into human decision-making cycles that is composed of three hierarchical levels: (level 1) the perception of the elements in the environment (level 2), the comprehension of their meaning, and (level 3) the projection of their future status. In the current work we use SA as a measure of an individual's success at performing cognitive information fusion to comprehensively understand the current status of events transpiring on a simulated battlefield.

At the cognitive or nodal level, the relationship between information, situation awareness, and task effectiveness has been extensively investigated in a number of ways including carefully controlled laboratory behavioral experiments. For example, Gonzalez and Wimisberg (2007) demonstrated that practice effectively improved information processing, the attainment of SA, and performance on dynamic decision tasks. Further, training reduced the relationship between individual cognitive abilities and SA to suggest that the cognitive demands of maintaining SA are reduced with practice. Also, a recent laboratory-based study examining human performance on simulated command and control tasks found that, contrary to expectations, increasing the volume of task-relevant information did not improve task performance, but instead reduced self-reported SA, leading to poorer task performance (Marusich et al., 2016). These results suggest that increasing the volume of information, even when it is accurate and task-relevant, is not necessarily beneficial to decision-making performance and may be detrimental to SA among team members. Military operations, however, are inherently complex human endeavors involving macro-cognitive processes that cannot be fully recreated or studied in the laboratory (Klein et al., 2003). As such, it is unclear whether these laboratory findings regarding the effects of practice and increasing volumes of information on SA also manifest themselves in naturalistic settings. As commanders and their staffs collectively face difficult, stressful, and dynamic challenges in managing battlefield operations, we need to determine the effects of information sharing, cognition, and training on their SA in more complex, real-world environments.

Warfare is chaotic and extremely complicated. Resolving the attendant ambiguity on the battlefield is both a cognitive and collaborative challenge of the first order. In these situations, human integration of networked information among the mission command staff is essential to successful military operations. A possible way to reduce the potentially detrimental effects of information overload is to distribute information processing tasks across the network—allowing separate people to process and act upon different sets of information (see, Kozlowski et al., 1999; Salas et al., 2008). In this case, a broad distribution of information and SA is essential for NEO. However, such distribution may also create added communication and coordination costs as well as additional dependencies, requiring each person in the network to maintain awareness of the dynamic situation and rely on the performance of others. Some research in military-relevant field exercises demonstrates a significant relationship between SA and the participants' awareness of the information in the central nodes of a team (Saner et al., 2009). This result suggests that SA is centralized and not broadly disseminated across the networked organization and that a person's role and position within an organization affects and potentially limits the level of shared SA that can be achieved. Our study scales up the results of these studies at the level of the individual and small teams to examine organizational network levels of performance.

The focus of our research is to examine and characterize the relationship between information sharing behavior and the distribution of SA in a real-world networked military organization. We examine how collaboration and information sharing among a large, networked mission command staff affects the attainment and distribution of individual SA across a 2-week real-world military exercise. Specifically, we construct network graphs from the record of staff communications throughout the exercise, and assess how the structure of these graphs relates to the SA of individuals within the network, as well as how this relationship evolves over the course of the exercise. Our results characterize the relations between volume of information, SA and performance and have major implications for training and systems design in NEO domains. Next, we describe this training event and our data collection and analysis.

## MISSION COMMAND TRAINING EXERCISE EVENT

The Mission Command Battle Laboratory at Fort Leavenworth, Kansas conducted a training event exercise focused primarily on the mission command operations of staff composed of a Division headquarters ($n = 46$) and two subordinate Brigade headquarters ($n = 21$, $n = 23$). Additional units and staff at echelons above and below the Division and Brigades participated in the training event exercise, with the size of the networked organization in excess of 200 ($n = 213$). The network architecture and digitized nature of the event allowed examination of staff communications in a distributed, network-enabled environment. Below we describe the defining characteristics of this military organization, and the nature of the tasks they were required to complete.

## Defining Characteristics of the Military Organization

The participants were active duty (and in a few cases retired): Soldiers and officers with operational staff experience who were assigned to differentiated, well-specified, and inter-dependent roles. Several staffs at different echelons participated, including a functional slide of a Division operations center and the staffs of a U.S. Heavy Brigade Combat Teams (mechanized) and a U.K. Coalition Brigade Combat Team. The units operated in a distributed fashion (U.S. units at Fort Leavenworth and the U.K. unit at the Land Warfare Centre in Warminster) over a communication network using specialized military command and control hardware and software. Within each unit, staff members carried out the duties of nine different functional cells. These cells included Command, Maneuver, Intelligence, Fires, Civil Affairs, Signal, Sustainment, Protection, and Liaison. Individual responses and responsibilities to a given scenario event in the training exercise depended upon adherence to established workflows and standard operating procedures both within the unit and functional cell.

Several additional small units and staffs were presented in the exercise, including high command elements of an International Joint Command, as well as a Civil Military Operations Center to facilitate coordination of joint, interagency (e.g., Department of State, United States Agency for International Development), intergovernmental, and multinational efforts. In addition, a third Infantry Brigade Combat Team was notionally represented; however, their area of operations was quiet and not fully exercised by scenario events. At the lowest level, a number of key role positions were staffed to represent Battalion level units in Army Aviation, Engineering, Military Police, and Sustainment (i.e., Counter Improvised Explosive Device). We used an electronic survey instrument to collect SA information from the Division, Heavy Brigade Combat Team, and the Coalition Brigade Combat Team, as these groups and their interoperability were the primary focus of the exercise. The high command elements, the Infantry Brigade Combat Team, and Battalion level units did not receive the electronic survey.

The military organization was staffed and convened specifically to execute and accomplish a particular 2-week long training mission. They worked interdependently and engaged in collaborative decision-making for mission planning and execution. The organization functioned as a *purposive social system*, where members are readily identifiable to each other by role and work interdependently to accomplish one or more collective objectives (Hackman and Katz, 2010). The responsibility for performing the various tasks and sub-tasks necessary for mission success is divided and assigned among the staff.

## Defining Characteristics of the Tasks

The training scenario in a military exercise generates many overlapping series of event-driven tasks, the resolution of which requires a high degree of coordination among the participating command and control staff. Researchers have long pointed out that the nature of a task has a great influence on the steps and processes a group uses to perform the work (e.g., Roby and Lanzetta, 1958; McGrath and Kravitz, 1982). The tasks of groups in the military domain considered here have four distinguishing features:
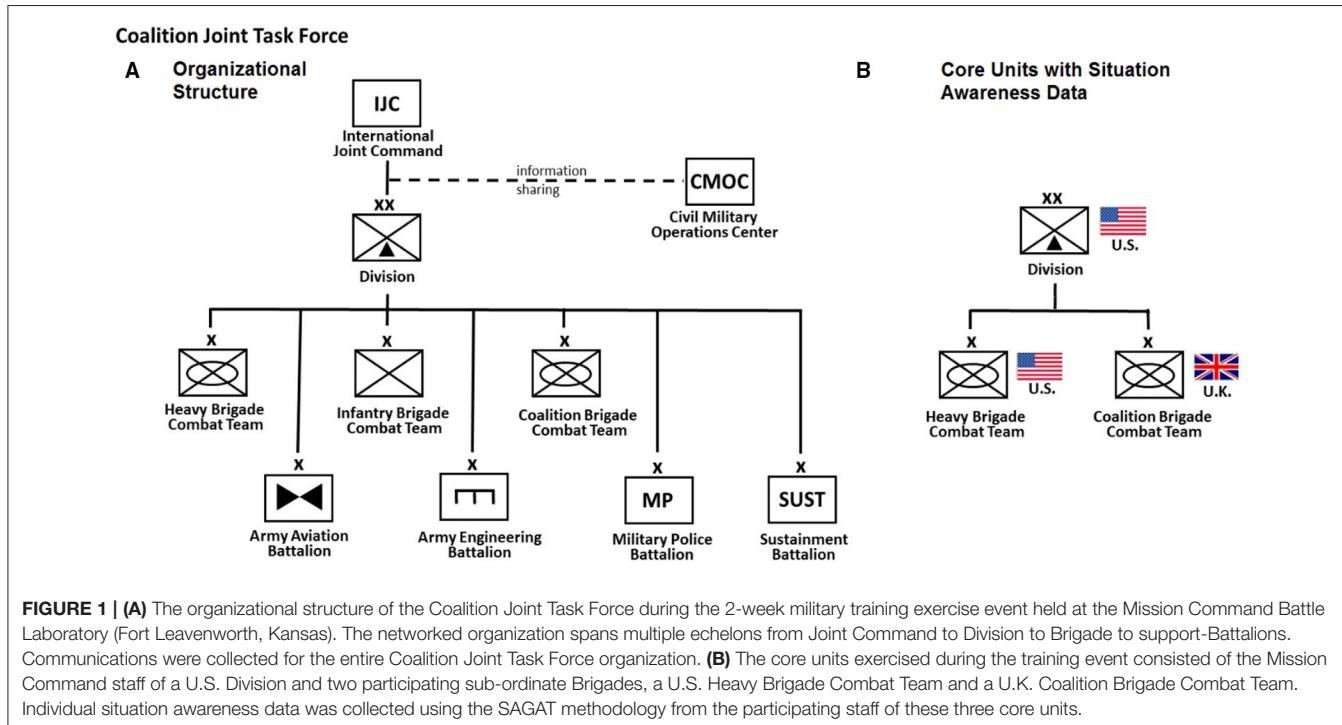
1. *Specific Presenting Problems:* The military command and control staff is tasked with addressing specific problems that occur in the unit's area of operations. The military staff organization must monitor key events and successfully plan and coordinate an effective response, given limited resources. The presenting problems may be kinetic events, such as responding to an improvised explosive device, or civil-military in nature, such as responding to a civil demonstration or safeguarding polling sites and maintaining a chain of custody in the transfer of voting ballots. At other times, the presenting problem may be a time-sensitive intelligence report of enemy activity that needs to be analyzed and corroborated. At any given time, the organization must coordinate a response to many such presenting problems.

2. *Adherence to specific tactics, techniques, and procedures:* The groups adhere to formalized military work routines and processes that are known in advance and involve delegation of specific work responsibilities to various sub-groups and individual staff members.

3. *Addressed immediately:* The group operates in an urgent, time-sensitive work environment and is required to immediately coordinate responses to work events that may have adverse cascading effects if not addressed in a timely manner.

4. *Results in collaborative work products that need to be coordinated and disseminated:* The group is expected to construct specific, detailed material products that will exist independently of the group process or the individual members themselves. For instance, the Commander and his command elements require regular reports from the staff in order to achieve situational awareness of the battlefield environment. The work processes themselves and the dissemination of both intermediate and final work products occur across the communication network as observable behaviors over time.

## Data Collected

### Communications Network

Telephone and email were two primary methods of direct communication between staff members during the exercise. For each email message sent and phone call made in our dataset, three pieces of information were automatically logged electronically: the sender, the receiver, and the time of the communication's initiation. The resulting full communications network consisted of: (a) an email network of 213 mission command staff members and 19168 correspondences, and (b) a telephone network of 3191 calls between 132 mission command staff members. The survey methodology, however, was only applied to the core units of the Coalition Joint Task Force organization. Thus, a subset of the email communications network (see **Figure 1**, right panel) is subsequently visualized and used for our statistical model

**FIGURE 1 | (A)** The organizational structure of the Coalition Joint Task Force during the 2-week military training exercise event held at the Mission Command Battle Laboratory (Fort Leavenworth, Kansas). The networked organization spans multiple echelons from Joint Command to Division to Brigade to support-Battalions. Communications were collected for the entire Coalition Joint Task Force organization. **(B)** The core units exercised during the training event consisted of the Mission Command staff of a U.S. Division and two participating sub-ordinate Brigades, a U.S. Heavy Brigade Combat Team and a U.K. Coalition Brigade Combat Team. Individual situation awareness data was collected using the SAGAT methodology from the participating staff of these three core units.

analysis of information and situation awareness. The telephone network was sparse and did not fully represent all the members of the core staff and thus not subjected to statistical model analysis.

## Situation Awareness Global Assessment Technique (SAGAT)

A valid and reliable method for assessing SA is essential for understanding whether information sharing behavior improves the SA of the personnel involved in the networked organization. Techniques such as the Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1995) and the Situation Awareness Rating Technique (SART; Taylor, 1990) have been applied in a number of organizational settings including military operations (Salmon et al., 2006), medical care environments (Wright et al., 2004), robot control (Chen et al., 2011), and industrial processes (Patrick et al., 2007).

In our electronic survey, we used the SAGAT, a widely used and validated SA measure (Endsley, 2000; Sonnenwald and Pierce, 2000) that makes use of a *pop quiz* memory probe technique to immediately present a set of questions to an individual regarding the state of their current task environment. The SAGAT methodology freezes the event to assess individual SA using targeted sets of online queries (multi-item quiz). The SAGAT methodology was developed and administered twice daily using online queries; at two predetermined times each day, an electronic questionnaire popped up on the computer monitor of each Mission Command staff member. After completing the questionnaire and submitting their responses, the Mission Command training exercise resumed.

Implementing the SAGAT requires advanced knowledge of the events so that a targeted set of queries can be developed

and administered to the participating Mission Command staff. Each set of SA questions was determined in consultation with the lead mission planner coordinating the training exercise event, who determined the best times to administer the SA queries. Significant mission events that were expected to occur prior to the query time were identified and questions that would assess SA on these relevant events were selected. The questions were developed from an SA requirements analysis conducted for various Army Mission Command staff positions using goal-directed task analysis methodology (see Bolstad and Endsley, 2003). During the event, subject-matter experts tailored the queries to the unfolding events and relevant aspects of mission in the area of operations for each Unit: US Division, UK Brigade, and US Brigade. The SA queries were broadly applicable, and not tailored to each role. Everyone received the same SAGAT queries but the answers were unit-specific. For example, the answers to the query "In your sector, which of the following CIVILIAN ACTIVITIES are currently occurring?" would be different for the US Brigade and the UK Brigade based on what was happening in their area of operations. Ground truth was determined based on tracking events in the simulation and feedback from subject-matter experts controlling the scenario-based exercise.

Each individual SA questionnaire included on average eight items from a total pool of 33 general queries. Unanswered questions were scored as incorrect. Questions were scored based on the participant's base unit. The data was collected by a contracted partner, SA Technologies Inc., to the Mission Command Battle Laboratory and provided to us in the aggregate for week 1 and week 2 of the exercise event. A sample set of queries is given in **Table 1**.

**TABLE 1 | Sample 19-item quiz administered to mission command unit using SAGAT methodology.**

1.   At this time, the MOST significant CIVILIAN event involves which of the following?

2.   At what LEVEL have CYBER ATTACKS been directed against Coalition operations in the last 4 h?

3.   Do you currently have troops in contact?

4.   Has a Commander's Critical Information Requirement been reported in the LAST 4 h?

5.   Have you received ACTIONABLE INTEL in the last 4 h REGARDING High Value Targets in your Area of Operations?

6.   How LONG has it been since the last MEDEVAC in your Area of Operations?

7.   In which portion of the Area of Operations was the LAST CALL for FIRES?

8.   In your sector, which of the following CIVILIAN ACTIVITES are currently occurring?

9.   The LAST REQUEST in your Area of Operations from CIVILIAN leaders was for which of the following?

10.  The MOST RECENT DETAINEES in your sector were engaged in which of the following BEFORE CAPTURE?

11.  What is the NATURE of the most recent REQUEST from COALITION/HOST nation partners?

12.  What type of targets will Counter Coalition Forces attack within the NEXT 2 h?

13.  What was the COALITION RESPONSE to the last attack in your sector?

14.  What was the NATURE of the last incident reported?

15.  What WEAPONS did the Counter Coalition Forces employ in the LAST attack in YOUR SECTOR?

16.  Which of the following best describes the TARGET of the last Counter Coalition Forces attack in your sector?

17.  Which of the following describe the OUTCOME of the last attack in your sector?

18.  Which of the following have been INCORRECT or EXAGGERATED in media reports in the last 4 h?

19.  Which of these INFRASTRUCTURE SERVICES are disrupted in your Area of Operations?

## RESULTS

### Social Network Visualization
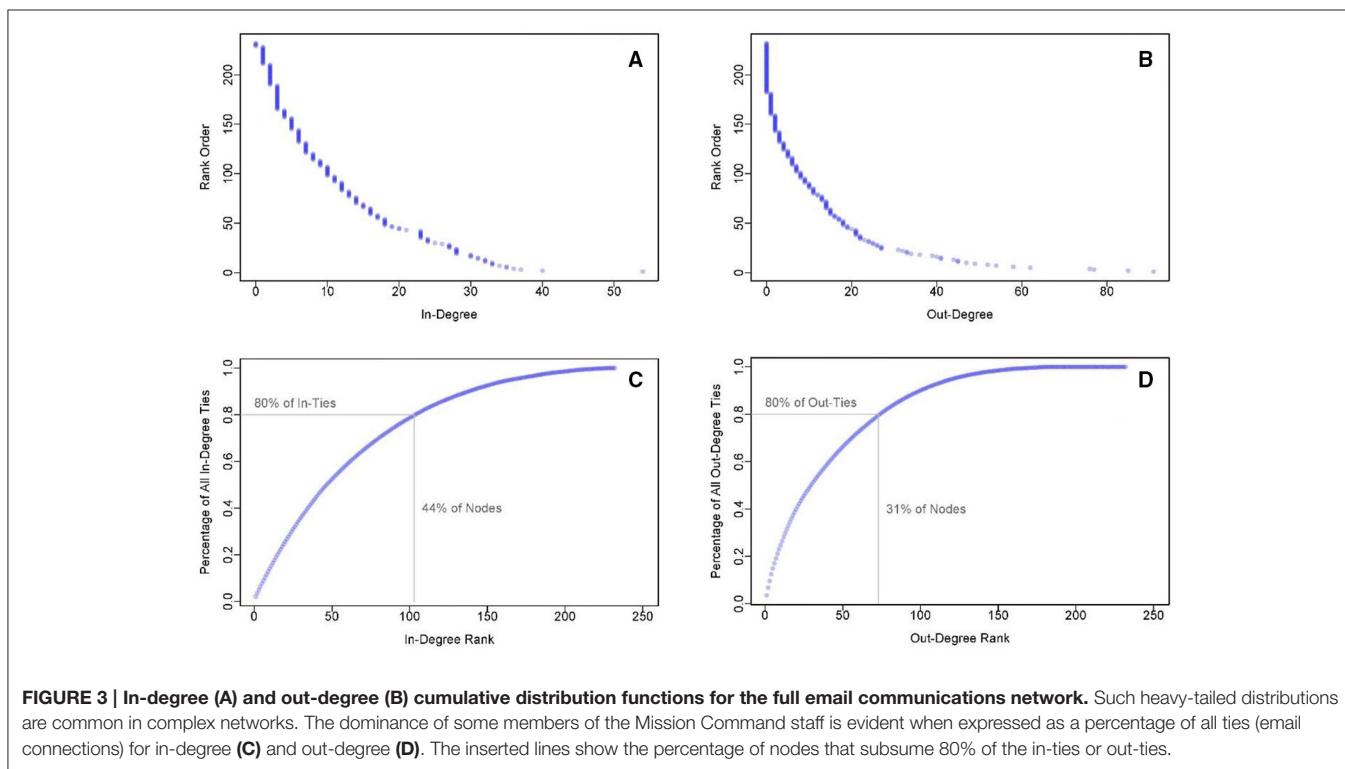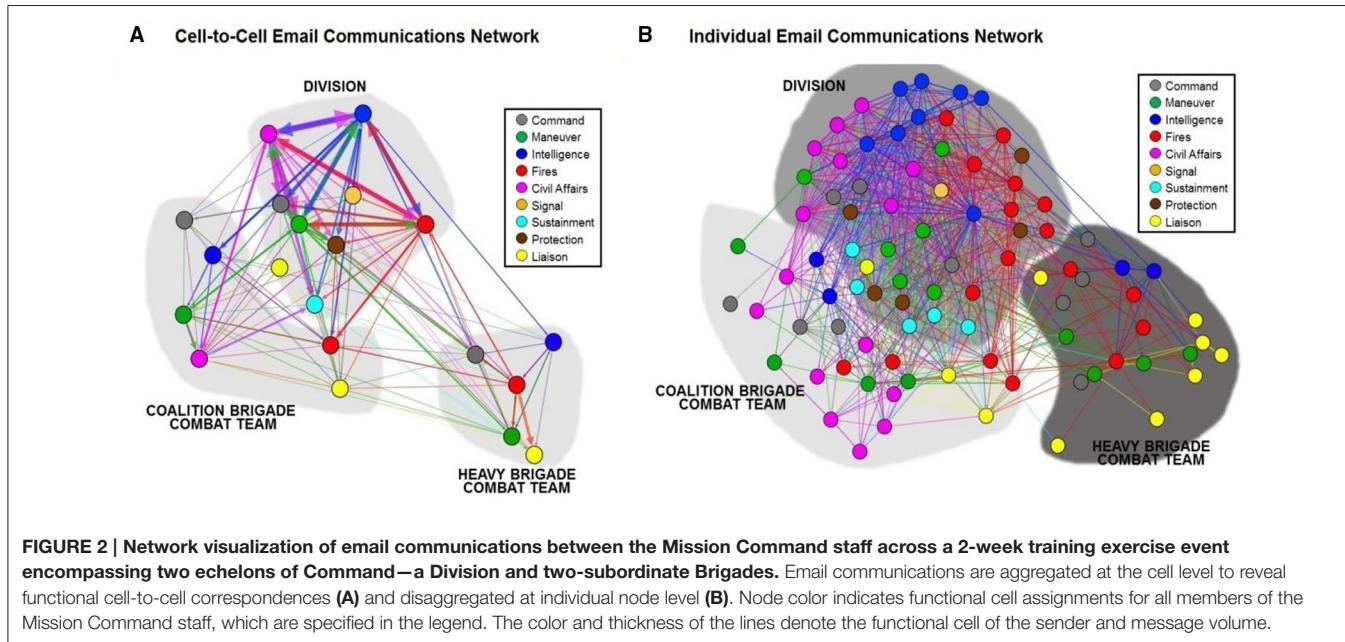
A network is defined as a set of nodes and the connections between them, called edges (undirected) or arcs (directed). In our organizational network of military command staff, the social collaborations are represented as directed email connections between individual nodes. The strength of a connection—number of email correspondences between nodes— is represented by the thickness of the line. At aggregate levels of analysis, the nodes can be grouped into units and cells to understand functional information flows. There were 45 individuals in Division roles, 23 in U.K. brigade roles, and 21 in U.S. Brigade roles, for a total of 89 nodes that were used in both our network visualization and subsequent statistical model and analysis. The pattern of email communications highlights the complex interdependencies and information sharing among the Mission Command staff (**Figure 2B**) and the diverse information flows between functional cells. The layout of the network visualization was produced using Gephi—an open-source network analysis and visualization software package—and is energized to minimize the overall variation in line length

using the Force Atlas (Jacomy, 2009) algorithm. This algorithm effectively centralizes the most highly-connected nodes and pushes the least connected nodes to the periphery. Our levels of analysis extend from the unit-level (e.g., Division, Brigade, and Battalion) to function-cell (Command, Maneuver, Intelligence, etc.) all the way down to characterizing individual staff members. The network visualization highlights the sheer complexity of current information sharing environments to achieve coordination and unity of effort among the Mission Command staff.

### Imbalanced Information Sharing

The distribution of email communications among the command staff is represented by in-degree and out-degree. The in-degree of a node is the number of individuals who send messages to that node. Conversely, the out-degree is the number of individuals who receive messages from that node. In our observed mission command network, a fundamental asymmetry exists in the degree that distribution of information flows among staff email collaborations. A few key individuals dominate information sharing among the staff. This is apparent in the cumulative distributions of in-degree and out-degree of email correspondences (see **Figures 3A,B**). These plots show the number of individuals with degree greater than or equal to a specified value. Most individuals in the network have very few connections, but a small number of individuals have many connections. Steeper drop-offs in these plots correspond to greater asymmetry in the degree distribution. The dominance of key members of the Mission Command staff conforms to a general network property of complex systems. The degree distributions of real-world networks are typically skewed and non-normal (i.e., non-Gaussian) with heavy tails (Barabási et al., 1999; Strogatz, 2001). Heavy-tailed distributions are so pervasive in real-world networks—turning up again and again in a wide variety of both natural and social phenomena, from earthquakes and floods to wealth, talent, and Internet behavior (West, 2012) that in organizational settings this phenomenon is known as Pareto's Law of the vital few (20%) and the trivial many (80%). At the macro level, Pareto (1897) first described imbalances in the wealth distribution of western countries such that 20% of the people owned 80% of the wealth. The seminal importance of this pioneering work is noted by West (2012, p. 78), who describes Pareto as "the first to have the modern vision of society as a network of reciprocal and mutually interdependent entities."

In our email communication network, key individuals at the tail of the degree distribution were found to dominate collaborations. The steeper drop-off of **Figure 3B**, as well as the more extended tail indicates that this was even more evident in the out-degree distribution than the in-degree distribution. We examine these degree imbalances in terms of the Pareto phenomenon (see **Figures 3C,D**). Degree rank is plotted on the x-axis, with 1 being the individual with the highest degree, and the percentage of all in-degree connections in the network belonging to that individual is plotted on the y-axis. Here again, a steeper curve indicates a greater imbalance in the degree distribution. We mark on the curves the points denoting how many individuals are responsible for

**FIGURE 2 | Network visualization of email communications between the Mission Command staff across a 2-week training exercise event encompassing two echelons of Command—a Division and two-subordinate Brigades.** Email communications are aggregated at the cell level to reveal functional cell-to-cell correspondences **(A)** and disaggregated at individual node level **(B)**. Node color indicates functional cell assignments for all members of the Mission Command staff, which are specified in the legend. The color and thickness of the lines denote the functional cell of the sender and message volume.



**FIGURE 3 | In-degree (A) and out-degree (B) cumulative distribution functions for the full email communications network.** Such heavy-tailed distributions are common in complex networks. The dominance of some members of the Mission Command staff is evident when expressed as a percentage of all ties (email connections) for in-degree **(C)** and out-degree **(D)**. The inserted lines show the percentage of nodes that subsume 80% of the in-ties or out-ties.

80% of the ties. In the in-degree distribution, 44% of the staff were responsible for 80% of the in-ties. In the out-degree distribution, only 31% of the nodes were responsible for 80% of the out-ties, nearing the classic Pareto distribution. Ultimately, this is interpretable as the implicit imbalance and pervasiveness of heavy-tailed distributions in complex networks.

## Exponential Random Graph Statistical Model

Exponential-family random graph models (ERGMs) are a family of statistical models widely used for inferential analysis of social network data (Hunter et al., 2008). Observed networks are standalone instances of many possible realizations of a given network. To support statistical inference about the structure
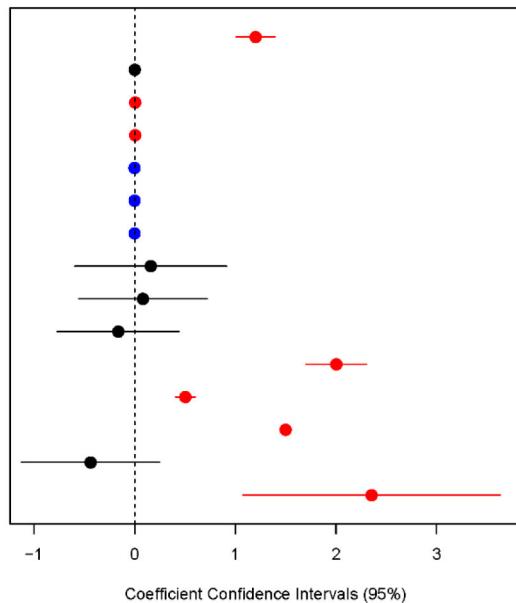
of a given network, an ERGM compares the similarity of the current observed network to the set of all possible alternative configurations. This allows us to establish a statistical baseline to infer the likelihood that the network could have expressed the observed structural characteristics at random. The ERGM models described below give the probability of observing a particular structural edge—an email connection—as a function of the model parameters, which are based on a variety of statistics from the network. The coefficients are not unlike those in a logistic regression, and can be interpreted as their effect on the log-odds of observing a given edge. In the email network, for example, the log-odds of observing an edge that reciprocates another edge is significantly higher than observing an edge that does not reciprocate an edge.

Using the ergm package in R (Handcock et al., 2016), we fit separate ERGM models to Week 1 and Week 2 of the exercise (Appendix). The model coefficients for each week are plotted in **Figure 4**. Results that are positive and statistically significant are colored red, results that are negative and significant are colored blue, and results that are not statistically significant are shaded black. The circle represents the value of the coefficient and the lines represent the accompanying 95% confidence interval. Due to the sheer number of communications in our dataset, some model coefficients have very small but significant effects even though they appear to sit on the 0 mark. We describe the effective terms of the statistical model in detail below.

## Robust Information Sharing Environment

Across both weeks, we find strong positive effects for within-cell homophily, reciprocity, triadic closure, and indegree. Homophily refers to the observation that networks often foster connections based on similarity (McPherson et al., 2001); in our case, defined as other individuals of the same functional cell (see **Figure 2**). These functional cells are well-defined and known according to the general staff system (Department of the Army Headquarters, 2015) and include: command, maneuver, intelligence, fires, civil affairs, signal, sustainment, protection, and liaison. During both weeks the model demonstrates a propensity for within-cell ties in the communication network. Reciprocity in directed email communications between two individuals (dyads) refers to the likelihood of mutual connections or email exchanges between them. We found a high propensity for reciprocity of email exchanges between individuals. That is, in a directed graph, if individual A sends email to individual B there is a strong likelihood that B also sends an email to A. More elaborate social structures arise when considering three individuals (triads) since a much wider set of interactions is possible among them. Triadic closure refers to a property of social networks that if relations exist between two pairs of individuals (A-B and A-C), then there is a strong likelihood of a tie (B-C) that completes the triangle of relations. Both reciprocity and triadic closure are common features of social networks (Scott, 2012). The model terms indegree, outdegree, and triadic closure were geometrically weighted to control for preferential attachment effects so that
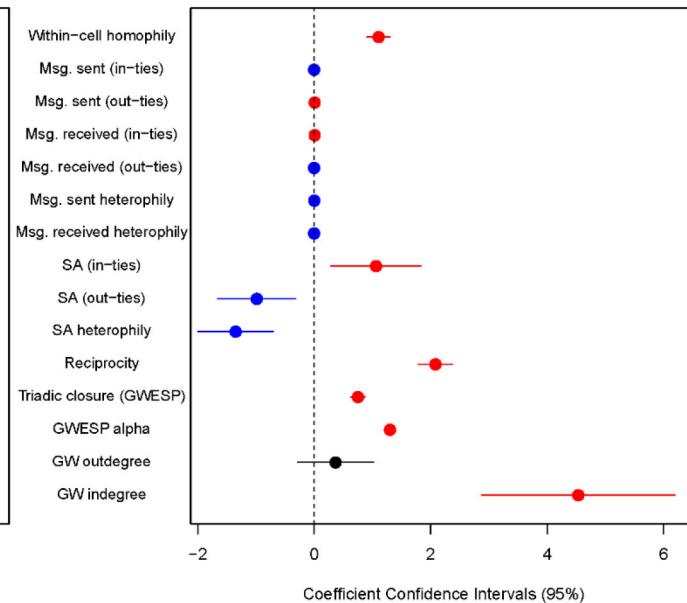


**FIGURE 4 | Exponential random graph statistical models of the email communication network during week 1 (left panel) and week 2 (right panel) of the Mission Command training event exercise.** The models describe the probability of observing any given edge as a function of the coefficients (log odds) in the statistical model. Results that are positive and significant are colored red, results that are negative and significant are colored blue, and results that are not statistically significant are colored black. The circle represents the statistical coefficient while the lines represent the 95% confidence interval for the coefficient. Note that given the large volume of messages some nodes have very small and significant effects even though they appear to be sitting on the 0 mark.

each additional shared partner has a declining positive impact on the probability of one or two persons forming a tie. This has been shown to work well in overcoming model degeneracy effects (i.e., bimodality) and in producing generalizable models that accommodate source and sink effects (see Hunter and Handcock, 2006; Hunter, 2007).

Our model also examines the association between tie formation and the number of messages sent or received, independent of degree. Across both weeks we find a positive, significant effect for the number of sent messages and out-degree [Msg. sent (out-ties)] and also between the number of received messages and in-degree [Msg. received (in-ties)]. As the volume of messages sent or received increases, so does the number of channels through which the individual sends or receives those messages. That is, rather than continuing to direct messages to a single partner or small set of partners, an individual who sends many messages is likely to send those messages to a large population of alters. The same is true for incoming messages. For a separate treatment of this dataset, see Marusich and Buchler (2016) for a detailed account and model of the overall email communication time series (by volume) and how it relates to an external work variable—the occurrence of significant simulated scenario events during the training event exercise.

## Information Sharing and Decreased Situation Awareness

Our central hypothesis is based on the tenets of NEO. Alberts and Garstka (2004) posit that increased information sharing in an organization improves situation awareness. The model coefficients for a link between situation awareness and in-degree [SA (in-ties)] or out-degree [SA (out-ties)] examine whether nodes with higher or lower situation awareness are more or less likely to send or receive ties. For Week 1, we obtained null results: the statistical coefficients were not significant as tie formation (number of in-ties and out-ties) was not associated with higher (or lower) levels of situation awareness among the mission command staff. In Week 2, however, we do find a relationship between SA and the propensity to receive and form network ties. Higher SA was associated with a higher tendency to form in-ties and a lower tendency to form out-ties. This challenges the hypothesis that increased information sharing improves situation awareness. The model coefficient for email in-ties and situation awareness was positive and significant, which indicates that those with high SA were more likely to be the recipients of ties. Receiving email (in-ties) implies a requirement for information, suggesting involvement in an organizational work process. On the other hand, sending email (out-ties) can be more material as it more directly advances an organizational work process especially if the information is processed and enhanced (value-added) and not just passed along. In other words, sending email is by definition an active process whereas receiving email is a passive process. Situation awareness is usually associated with an active process of constructing a mental model of the current events (Endsley, 2000), and thus should be associated with active work processes such as processing and sending email.

A plausible explanation of these results is that lower situation awareness is associated with work demands. The implication is that sending email demands attentional resources from the user and thus detracts from their overall situation awareness due to multi-tasking demands—in the same way that chatting with a passenger might distract a driver from paying attention to the route. An alternative explanation is that processing and sending email is associated with addressing specific challenges and fashioning work products, so that attention is not broadly allocated and instead tightly constrained and focused intently on processing a subset of features and events in the battlespace. A plausible explanation for the in-tie effect is that people with greater knowledge are likely to be tapped as potential sources of information and expertise, per transactive memory theory (Contractor and Monge, 2002; Borgatti and Cross, 2003), a rich-get-richer effect. Broadening the number of email out-ties was associated with lower situation awareness, perhaps because of the complexity of the operational environment outpaces human cognitive capabilities. In our broadly collaborative and information-rich Mission Command network, the accumulation of information can occur quite rapidly. In such cases, it can be difficult and time-consuming for the human operator to process relevant information and support workflows due to overwhelming volume of information and variety of different sources (email, chatrooms, maps with graphical overlays, imagery, video). An alternative explanation may be that those with higher SA did not find it necessary to reach out to others to obtain mission-critical information, with more in-ties they already had a firm grasp of their operating environment. Given the limits of causal inference, it is also possible that individuals with lower SA may send out more requests for information; disambiguating the direction of the effect requires an analysis of email content, which we do not have.

We note the development of an additional communication pattern that was cemented by week 2, according to the model coefficients. During the second week, we found positive effects for sending messages and establishing additional ties, in addition to receiving messages and receiving additional ties. We found that individuals with more incoming ties were less likely to *send* larger volumes of emails, while those with more outgoing ties less likely to *receive* more messages. Using the standard terminology of the network information flow perspective (Zachary, 1977; Ahlswede et al., 2000), taken together these four effects suggest that certain individuals increasingly act as *sources* or *sinks* of information in the networked organization. That is, they acted as either broadcasters or attractors of information. This reinforces the primacy of our earlier result that information is not shared equally in the network with Pareto-type imbalances to the in-degree and out-degree distributions.

## Homophily in Email Communications

Homophily effects offer further insight into the pattern of SA that we observe in the network. During the second week we find a significant, negative effect for SA heterophily. That is, individuals with larger differences in SA are less likely to form ties with one another (i.e., lower SA individuals tend to communicate among themselves and higher SA individuals tend to communicate among themselves). This is another emergent property of the network, as we did not observe this pattern during

week 1. Over time the network appears to become stratified with respect to SA. This could be the result of deliberate action— individuals with higher SA may reach out to others with high SA while avoiding those with lower SA—or this may be an outcome of the structural configuration of the network—those with access to information that enhances SA were unable to diffuse that information to parts of the network and, as a result, SA declined among those subgroups. The stratification of SA in this network is problematic for organizational performance and this problem deserves further attention both to advance organizational research and improve the effectiveness of military performance during training exercises. One possibility is that the organization is essentially divided into information processors whose job it is to understand the situation and who receive and send an inordinate number of email messages (and have high SA as a result) and other members of the staff whose job is much more delimited and circumscribed to particular tasks and thus send and receive fewer email messages (and have low SA as a result). In other words, it is possible that the pattern of communications reflects a division of labor that emerged among the mission command staff, as their functional role assignments were fixed and relate to their chosen military occupational specialty, an enduring property of their profession.

## DISCUSSION

At a large-scale, 2-week military training event exercise, we conducted a social network analysis of email communications among a multi-echelon mission command staff to assess the commonly held assumption that increased information sharing improves situation awareness among the staff in complex networked operational environments. Results from our exponential random graph models challenge this assumption, as we found that increased email output was associated with lower individual situation awareness. Conversely, higher SA was associated with a lower probability of out-ties, so that sending too many messages broadly to other individuals decreased the likelihood of attaining SA. This challenges our hypothesis that increased information sharing improves situation awareness and also supports a recent laboratory studies that increasing task-relevant information did not improve task performance, but instead reduced self-reported SA, leading to poorer task performance (Marusich et al., 2016). In addition, we observed two strong effects of homophily in email communication. Links were more commonly formed between members of the command staff with similar functions and levels of situation awareness, than between two individuals with dissimilar functions and levels of situation awareness. These findings have major implications that challenge the current conceptual framework of NEO (Alberts and Garstka, 2004) which posits that robust networking and information sharing act as a positive feedback loop resulting in greater situation awareness and mission effectiveness in military operations. These and other results highlight several major growing pains for networked organizations and military organizations in particular.

## Unequal Information Sharing

The first growing pain for organizations is that information is not shared equally, even in robust and relatively unconstrained information sharing environments. In our observed mission command network, there were large imbalances to information sharing as a few key individuals dominated information sharing among the staff. Most individuals in the network have very few email connections, but a small number of individuals have very many connections. The dominance of key members of the Mission Command staff conforms to a general Pareto-type network property of complex systems (see West, 2012). At network levels of interaction, understanding the social and cognitive dynamics that give rise to Pareto's law constitutes a fundamental question for network science research. Intuitively, it is possible that the degree distribution imbalance occurs whenever there is a fundamental imbalance in the value of individuals in the network. In our mission command network the value of individuals is reflected by military rank/experience and the primacy of certain functional role-positions. If so, this phenomenon could extend beyond military networks to include any workgroup structured using an organizational hierarchy, especially corporations, bureaucracies, departments, and work-groups among others.

In networked organizations, the sheer volume and rapid pace of information and communications received and readily accessible through computer networks can be overwhelming to individuals, resulting in data overload from diverse data sources, multiple data formats, and large data volumes. The need to integrate and interpret information in massive data environments and the macrocognitive processes involved in fashioning a coherent understanding is commonly referred to as sensemaking (Klein et al., 2006). Given the Pareto-type imbalances to the email degree distributions, it is likely that some individuals in the network are beyond their functional cognitive capacity to process and make sense of so much information. It is the case that in complex tasks, limitations in cognitive resources and processes have been shown to give rise to many cognitive biases that distort human decision making (Lebiere et al., 2013). However, humans are remarkably resilient in adapting to the complexity and functional limitations of their environment. Researchers have documented a variety of cognitive strategies and systematically examined the tradeoffs and shortcuts involved in overcoming fixed limits to human information processing capacities, such as attention bottlenecks and memory limitations (see Reitter and Lebiere, 2012). One of those tradeoffs and associated techniques is whether to share raw information, providing all the needed information at the cost of potentially overwhelming attentional demands, or high-level summaries and conclusions, requiring context-sensitive filtering and inference that may miss critical issues in the presence of information stovepipes (Tang et al., 2015).

As a practical consideration, following the business maxim put forward by Koch (2011) in his book, *The 80/20 Principle*, efforts should be made to support this vital 20% that also generates 80% of the work. This suggests that technological solutions and training regimens should focus on supporting the vital 20% of the networked staff driving most of the collaborations

(for a decision-support agent approach, see Buchler et al., 2014). The long-tailed distributions of communications have major implications for the psychological and social sciences as many parametric statistical approaches and human performance modeling tools assume some degree of normality in the processes they model (Warwick et al., 2013). As discussed below, understanding how cognition is manifest at network levels of interaction represents a challenge and opportunity for macro-cognitive researchers.

Scientists and engineers have developed many approaches for understanding and predicting individual and group state, behavior, cognition, and performance in the context of teams, organizations, and societies; with each approach being limited in resolution, validity, and insight into the human condition. Understanding how humans interact and adapt within dynamic, complex, natural environments remains a pressing and challenging scientific problem. Recent technological advances have lead researchers and information technology firms (e.g., Navaroli and Smyth, 2015) to leverage vast quantities of data from various human, information, and communication networks to make interpretations and predictions about humans and the context in which they are operating. Network science approaches allow both the organizational context and real-world human behavior to be jointly analyzed and interpreted. However, network science focuses on the interactions between decision-makers and their emergent social phenomena, often oversimplifying many cognitive aspects of the individual nodes. This represents both a challenge and opportunity for macro-cognitive research to define cognitive processes that occurs at the "nodal level" in real-world contexts, such as decision-making under uncertainty and sense-making. In essence, the defining challenge is to understand the cognitive processes that give rise to the heavy-tailed statistics seen at network levels of interaction. For instance, a cognitive mechanism formally implemented at the nodal level as a priority communication model—sorting communication messages by importance in a queue (i.e., email inbox)—was shown in simulation to give rise to the patterns of real-world bursty communication timings observed at the network level (Vázquez et al., 2006).

## Organizational Stovepipes

A second growing pain for organizations is one of breaking open "information stovepipes" or existing socio-technical limitations that restrict the free flow of information and communications (e.g., Bateman, 1996). The flow of information among the Mission Command staff involves the timely push and pull of information and knowledge products to and from adjacent, higher, and lower functional cells and units. The distribution of information, however, was largely constrained to and adhered to unit structure of the organization, and thus largely occurred within functional cell assignments. The pattern of communications in our networked organization conform to well-established principles of social networks as we observed strong effects of reciprocity, triadic closure, and within-cell homophily that were governed by their functional cell assignment. These functional cells are well-defined and known according to the general staff system (Department of the Army Headquarters,

2015) and include: command, maneuver, intelligence, fires, civil affairs, signal, sustainment, protection, and liaison. This is the hallmark of a stovepiped organization where information is bottled-up and not widely shared among diverse individuals in the organization. The general pattern of results raise fundamental questions as to the macro-cognitive mechanisms existing at the individual node level that give rise to the patterns observed at the level of the networked organization.

It is not clear how to promote diverse heterophilous ties within an organization. Currently, two theories have been advanced for a lack of heterophilious ties in organizational settings (Chung et al., 2000). First, rank confers status within the Mission Command network and higher-status individuals and organizations in the multi-echelon hierarchy may see their status reduced by ties to lower-status individuals and organizations (Benjamin and Podolny, 1999). In this case, the propensity is to communicate with high-rank individuals. It is possible that this propensity to concentrate communications to high-ranking individuals can drive the types of in-degree imbalances we observed in our email communication network. Indeed, many of the individuals at the tail of the degree distribution are high-ranking principal members of the Division mission command staff. Second, individuals and organizations may have access to unequal information quality which reduces the value proposition of information exchanges between individuals with dissimilar situation awareness. In addition, maintaining heterophilous relationships across functional cells and across unit echelons can be, in practice, quite difficult due to dissimilar work processes, complex information requirements, lack of awareness, and the multitude of disparate information systems that can constrain such collaborations.

## Emergence of Information Sources and Sinks

A third growing pain is the emergence of individuals that function increasingly as sources and sinks of information in the networked organization. From an information flow perspective (Ahlswede et al., 2000), network ties are social channels that allow the flow of information throughout the organization. We observed that by Week 2 of the exercise, with more incoming ties individual members of the Mission Command staff were less likely to send out larger volumes of emails. With more outgoing ties, individuals were also less likely to receive more messages. This suggests that certain individuals increasingly act as sources and sinks in the networked organization and suggests a specialization of information sharing behavior as either broadcasters or attractors of email communications. This also reinforces the primacy of our earlier result that information is not shared equally in the network with Pareto-type imbalances to the in-degree and out-degree distributions. Furthermore, these source and sink effects are emergent properties of the organization. These results support earlier research from military field exercises demonstrating that SA is concentrated to a few select individuals and linked to the participants' awareness of the information in the central nodes of a team (Saner et al., 2009).

## Stratified Situation Awareness

A fourth growing pain was that over time our organizational network appears to become stratified with respect to SA—an effect of homophily with respect to SA. Those with high situation awareness were likely to have ties to others who also have high SA. Conversely, those who have low SA were likely to have ties to others who also have low SA, and thus have impoverished information flows. The effects of homophily and SA emerged during Week 2 of the military training event exercise and is likely a self-reinforcing phenomenon. This could be the result of deliberate action—individuals with higher SA may reach out to others with high SA while avoiding those with lower SA— or this may be an outcome of the structural configuration of the network—those with access to information that enhances SA were unable to diffuse that information to parts of the network and, as a result, SA declined among those subgroups.

The stratification of SA in this network is problematic for organizational performance and this problem deserves further attention both to advance organizational research and improve the effectiveness of military performance during training exercises. One possibility is that the organization is essentially divided into information processors whose job it is to understand the situation and who receive and send a lot of email messages (and have high SA as a result) and other members of the staff whose job is much more delimited and circumscribed to particular tasks and thus send and receive fewer email messages (and have low SA as a result). In other words, it is possible that the pattern of communications reflect a division of labor that emerged among the mission command staff, as their functional role assignments were fixed and relate to their chosen military occupational specialty, an enduring property of their profession.

Given our results, it is a likely that the stratification of SA emerges as a consequence of the information sharing behavior of the organization to include homophilous ties (and lack of heterophily) and Pareto-type imbalances in the degree distribution. An open question that can be tackled through simulation is whether one or more general mechanisms can produce the observed pattern of results as an emergent process of the organization. That is, it is possible that the generally observed properties of email homophily, reciprocity, and triadic closure can result in Pareto-type imbalances in the degree distribution, which can in turn lead to organizational stovepipes among the staff, sources and sink effects, and ultimately the stratification of situation awareness. Overall, our result suggests that SA is stratified across the networked organization and that a person's role and position within an organization affects and potentially limits the level of shared SA that can be achieved.

Our approach focused on relating individual SA to network levels of interaction among the Mission Command staff. A more nuanced approach for future research involves defining SA in relation to the information requirements required for a given staff role position and unit. Each member of the team provides valuable and critical information within and across roles. For instance, team members in different roles (e.g., commanders, intelligence officers, logistics officers) have common information requirements and also some that are unique to their functional role (Artman and Garbis, 1998). In this case, SA is defined at the aggregate team level and furthermore is also used to define

common or overlapping information requirements necessary for shared SA. Although potentially useful to support teammates, it is not necessary for each member of the team to have all the information needed by others on the team. It is important, however, that each team member understands what information is needed to support multiple role positions. Shared SA refers to the degree to which team members have the same SA on a defined set of shared information requirements (Endsley and Jones, 2013). For effective team performance, Team SA refers to the sum total of information and degree to which each team member obtains the SA needed to fulfill his or her responsibilities (Endsley, 1995). It is the case that these are overlapping and mutually defined sets of information that are derived from individual SA.

Many of these challenges faced by our Mission Command staff reflect broad trends and challenges in networked organizations and how to effectively manage the systematic convergence of people, information, and technology in work-directed networked organizations. It is likely that many of the findings that we observed in our Mission Command network are also evident in other organizations.

## CONCLUSION

The military transformation to NEO has proceeded under a conceptual framework that attempts to exploit the increasing interconnectedness between organizational units to allow more communication, information sharing, cooperation and thereby flexibility, adaptability, and mission effectiveness (Alberts, 2002; Alberts and Hayes, 2003). Our results highlight many challenges (i.e., growing pains) to NEO and the need for fundamental research to guide this transformation; much of the rapidly growing literatures in network science, organizational and team processes, and cognitive science do not fully address many of the presenting problems of complex operational environments, macro-cognition, human-in-the-loop systems, and the defining characteristics of work-driven organizations. The vast majority of insights have been gained through laboratory research using highly controlled contexts and environments. Many of these laboratory studies employ reductive scientific approaches (i.e., divide and conquer) that do not scale to complex real-world operations or larger networks and organizational settings. Recent advances in technology have led researchers and industry to leverage vast quantities of data from various human, information, and communication networks to make interpretations and predictions about humans and the context in which they are operating. Such "big science" approaches are fundamentally multi-disciplinary endeavors involving teams of scientists and engineers that embrace the complexity of real-world phenomena to examine network levels of interaction. Embracing complexity is a key challenge and conceptually is a paradigm-shift for science. Such "big science" approaches will certainly yield fundamental insights and understanding into many complex real-world phenomena, but may not be able to completely predict complex real-world phenomena that are non-deterministic, non-linear, and sensitive to initial conditions and feedback loops (see Arney et al., 2015).

## AUTHOR CONTRIBUTIONS

NB contributed to the ideas, design, execution of the study as well as the analyses of results and write up of the manuscript. SF contributed to data preparation, analyses, particularly social network analyses, and write up of results. LM contributed to data preparation, data analyses, and write up of methods and results. DU contributed to material preparation, data collection, and execution of the study. CL contributed to writing the manuscript, interpreting the results and framing the introduction. CG contributed to writing the manuscript, framing the introduction, and interpreting the results.

## ETHICS STATEMENT

This study was carried out in compliance of federal and Army Research Laboratory regulations requiring Institutional Review Board review of all research involving human subjects prior to the initiation of a research protocol to ensure the safe and ethical treatment of humans as subjects in research.

## ACKNOWLEDGMENTS

## REFERENCES

Ahlswede, R., Cai, N., Li, S. Y. R., and Yeung, R. W. (2000). Network information flow. *Inform. Theor. IEEE Trans.* 46, 1204–1216. doi: 10.1109/18.850663

Alberts, D. (2002). *Information Age Transformation: Getting to a 21st Century Military.* Washington, DC: Command and Control Research Program (CCRP) Publications.

Alberts, D., and Garstka, J. (2004). *Network Centric Operations Conceptual Framework Version 2.0.* Technical Report, US Office of Force Transformation and Office of the Assistant Secretary of Defense for Networks and Information Integration, US Department of Defense.

Alberts, D., Garstka, J., and Stein, F. (1999). *Network Centric Warfare: Developing and Leveraging Information Technology.* Washington, DC: CCRP Publication Series, Department of Defense.

Alberts, D. S., and Hayes, R. E. (2003). *Power to the Edge: Command and Control in the Information Age.* Washington, DC: Command and Control Research Program (CCRP) Publications.

Arney, C., Koban, D., Juliano, N., and Sobiesk, M. (2015). "Tomorrow's science connections: networks, measures, motivations," in *23rd ARL USMA Technical Symposium* (Aberdeen Proving Ground, MD).

Artman, H., and Garbis, C. (1998). "Team communication and coordination as distributed cognition," in *9th Conference of Cognitive Ergonomics* (Limerick), 151–156.

Barabási, A. L., Albert, R., and Jeong, H. (1999). Mean-field theory for scale-free random networks. *Phys. A Stat. Mech. Appl.* 272, 173–187. doi: 10.1016/S0378-4371(99)00291-5

Bateman, R. L. (1996). Force XXI and the death of *Auftragstaktik. Armor* 1, 13–15.

Bedny, G., and Meister, D. (1999). Theory of activity and situation awareness. *Int. J. Cogn. Ergon.* 3, 63–72. doi: 10.1207/s15327566ijce0301_5

Benjamin, B. A., and Podolny, J. M. (1999). Status, quality, and social order in the California wine industry. *Admin. Sci. Q.* 44, 563–589. doi: 10.2307/2666962

Blasch, E., Bossé, É., and Lambert, D. A. (2012). *High-Level Information Fusion Management and Systems Design.* Boston, MA: Artech House.

Blasch, E. P., Breton, R., and Valin, P. (2011). "Information fusion measures of effectiveness (MOE) for decision support," in *Proc. SPIE 8050, Signal Processing, Sensor Fusion, and Target Recognition XX, 805011*, ed Ivan Kadar (Orlando, Fl). doi: 10.1117/12.883988

Bolstad, C. A., and Endsley, M. R. (2003). "Measuring shared and team situation awareness in the army's future objective force," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Vol. 47 (Denver, CO: SAGE Publications), 369–373.

Borgatti, S. P., and Cross, R. (2003). A relational view of information seeking and learning in social networks. *Manage. Sci.* 49, 432–445. doi: 10.1287/mnsc.49.4.432.14428

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009). Network analysis in the social sciences. *Science* 323, 892–895. doi: 10.1126/science.1165821

Buchler, N., Marusich, L. R., and Sokoloff, S. (2014). "The Warfighter Associate: decision-support software agent for the management of intelligence, surveillance, and reconnaissance (ISR) assets," in *Proceedings of SPIE, Vol. 9079, 907902, International Society for Optics and Photonics, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR V, SPIE Defense, Security, and Sensing,* ed M. A. Kolodny (Baltimore, MD). doi: 10.1117/12.2054646

Castells, M. (2009). *The Rise of the Network Society: The Information Age: Economy, Society, and Culture,* Vol. 1. Malden, MA: Wiley-Blackwell.

Chen, J. Y., Barnes, M. J., and Harper-Sciarini, M. (2011). Supervisory control of multiple robots: human-performance issues and user-interface design. *IEEE Trans. Syst. Man Cybernet. C Appl. Rev.* 41, 435–454. doi: 10.1109/TSMCC.2010.2056682

Chung, S. A., Singh, H., and Lee, K. (2000). Complimentarity, status similarity and social capital as drivers of alliance formation. *Strateg. Manage. J.* 21, 1–22.

Contractor, N. S., and Monge, P. R. (2002). Managing knowledge networks. *Manage. Commun. Q.* 16:249. doi: 10.1177/089331802237238

Department of the Army Headquarters (2015). *Field Manual 6-0 Commander and Staff Organization and Operations.* Washington, DC.

Department of Homeland Security (2013). *DHS Information Sharing and Safeguarding Strategy.* Available online at: http://www.dhs.gov/sites/default/files/publications/12-4466-dhs-information-sharing-and-safeguarding-strategy-01-30-13--fina%20%20%20.pdf

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Hum. Factors J. Hum. Factors Ergon. Soc.* 37, 32–64. doi: 10.1518/001872095779049543

Endsley, M. R. (2000). "Theoretical underpinnings of situation awareness: a critical review", in *Situation Awareness Analysis and Measurement.* eds M. R. Endsley and D. J. Garland (Mahwah, NJ: Lawrence Erlbaum Associates), 3–32.

Endsley, M. R., and Jones, W. (2013). "Situation awareness," in *The Oxford Handbook of Cognitive Engineering,* eds J. D. Lee, and Alex Kirlik (New York, NY: Oxford University Press), 88–108.

Endsley, M. R., and Kiris, E. O. (1995). *Situation Awareness Global Assessment Technique (SAGAT) TRACON Air Traffic Control Version User Guide.* Lubbock, TX: Texas Tech University.

Endsley, M. R., and Smolensky, M. W. (1998). "Situation awareness in air traffic control: the picture," in *Human Factors in Air Traffic Control,* eds M. W. Smolensky and E. S. Stein (San Diego, CA: Academic Press), 115–154.

Federal Aviation Administration (2014). *NextGen Implementation Plan.* Washington, DC.

Gleick, J. (2012). *The Information: A History, a Theory, a Flood.* New York, NY: Random House LLC.

Gonzalez, C., and Wimisberg, J. (2007). Situation awareness in dynamic decision making: effects of practice and working memory. *J. Cogn. Eng. Decis. Making* 1, 56–74. doi: 10.1177/155534340700100103

Hackman, J. R., and Katz, N. (2010). Group behavior and performance. *Handb. Soc. Psychol.* 3:32, 1208–1251. doi: 10.1002/9780470561119.socpsy002032

Handcock, M., Hunter, D., Butts, C., Goodreau, S., Krivitsky, P., and Morris, M. (2016). *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks.* The Statnet Project (http://www.statnet.org). R package version 3.6.0. Available online at: http://CRAN.R-project.org/package=ergm

Hauss, Y., and Eyferth, K. (2003). Securing future ATM-concepts' safety by measuring situation awareness in ATC. *Aerosp. Sci. Technol.* 7, 417–427. doi: 10.1016/S1270-9638(02)00011-1

Hunter, D. R. (2007). Curved exponential family models for social networks. *Soc. Networks* 29, 216–230. doi: 10.1016/j.socnet.2006.08.005

Hunter, D. R., and Handcock, M. S. (2006). Inference in curved exponential family models for networks .*J. Comput. Graph. Stat.* 15, 565–583 doi: 10.1198/106186006X133069

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., and Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *J. Stat. Softw.* 24:nihpa54860. doi: 10.18637/jss.v024.i03

Jacomy, M. (2009). *Force-Atlas Graph Layout Algorithm.* Available online at: https://gephi.org/tutorials/gephi-tutorial-layouts.pdf

Kaber, D. B., Endsley, M. R., Wright, M. C., and Warren, H. (2002). *The Effects of Levels of Automation on Performance, Situation Awareness, and Workload in An Advanced Commercial Aircraft Flight Simulation.* Final Report: NASA Langley Research Center Grant# NAG-1-01002, Hampton, VA: NASA Langley Research Center.

Keller, J., Lebiere, C., Shay, C. R., and Latorella, K. (2004). "Cockpit system situational awareness modeling tool," in *Human Performance, Situation Awareness and Automation: Current Research and Trends: HPSAA II,* Vol. 2, eds D. A. Vincenzi, M. Mouloua, and P. A. Hancock (Daytona Beach, FL: Psychology Press), 66–71.

Klein, L. A. (2004). *Sensor and Data Fusion: A Tool For Information Assessment and Decision Making,* Vol. 324. Bellingham, WA: SPIE Press.

Klein, G., Moon, B., and Hoffman, R. R. (2006). Making sense of sensemaking 2: a macrocognitive model. *IEEE Intell. Syst.* 21, 88–92. doi: 10.1109/MIS.2006.100

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., and Hollnagel, E. (2003). Macrocognition. *IEEE Intell. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Koch, R. (2011). *The 80/20 Principle: The Secret to Achieving More with Less.* New York, NY: Crown Business.

Kozlowski, S. W., Gully, S. M., Nason, E. R., and Smith, E. M. (1999). "Developing adaptive teams: a theory of compilation and performance across levels and time,"in *The Changing Nature of Performance: Implications for Staffing, Motivation, and Development,* ed Pulakos (San Francisco, CA: Jossey-Bass Publishers), 240–292.

Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., et al. (2013). A functional model of sensemaking in a neurocognitive architecture. *Comput. Intell. Neurosci.* 2013:921695. doi: 10.1155/2013/921695

Marusich, L. M., and Buchler, N. (2016). Time series modeling of Army mission command communication networks: an event-driven analysis. *J. Comput. Math. Organ. Theor.* doi: 10.1007/s10588-016-9211-7. Available online at: http://link.springer.com/article/10.1007/s10588-016-9211-7

Marusich, L. R., Bakdash, J. Z., Onal, E., Michael, S. Y., Schaffer, J., O'Donovan, J., et al. (2016). Effects of information availability on command-and-control decision making performance, trust, and situation awareness. *Hum. Factors* 58, 301–321. doi: 10.1177/0018720815619515

Matthews, M. D., Strater, L. D., and Endsley, M. R. (2004). Situation awareness requirements for infantry platoon leaders. *Mil. Psychol.* 16, 149. doi: 10.1207/s15327876mp1603_1

McGrath, J. E., and Kravitz, D. A. (1982). Group research. *Annu. Rev. Psychol.* 33, 195–230. doi: 10.1146/annurev.ps.33.020182.001211

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 27, 415–444. doi: 10.1146/annurev.soc.27.1.415

Navaroli, N., and Smyth, P. (2015). "Modeling response time in digital human communication," in *Ninth International AAAI Conference on Web and Social Media.* Oxford, UK.

Pareto, V. (1897). *Cours d'Economie Politique.* Lausanne; Paris.

Patrick, J., James, N., and Ahmed, A. (2007). *Awareness of Control Room Teams.* Paris: Presses Universitaires de France.

Reitter, D., and Lebiere, C. (2012). "Social cognition: memory decay and adaptive information filtering for robust information maintenance," in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI*-12*).* Toronto, ON.

Roby, T. B., and Lanzetta, J. T. (1958). Considerations in the analysis of group tasks. *Psychol. Bull.* 55:88. doi: 10.1037/h0047233

Salas, E., Goodwin, G. F., and Burke, C. S. (2008). *Team Effectiveness in Complex Organizations: Cross-Disciplinary Perspectives and Approaches.* New York, NY: Routledge.

Salimi, N., and Vita, R. (2006). The biocurator: connecting and enhancing scientific data. *PLoS Comput. Biol.* 2:e125. doi: 10.1371/journal.pcbi.0020125

Salmon, P., Stanton, N., Walker, G., and Green, D. (2006). Situation awareness measurement: a review of applicability for C4i environments. *Appl. Ergon.* 37, 225–238. doi: 10.1016/j.apergo.2005.02.001

Saner, L. D., Bolstad, C. A., Gonzalez, C., and Cuevas, H. M. (2009). Measuring and predicting shared situation awareness in teams. *J. Cogn. Eng. Decis. Making* 3, 280–308. doi: 10.1518/155534309X474497

Scott, J. (2012). *Social Network Analysis.* Los Angeles, CA: Sage.

Snijders, C., Matzat, U., Reips, U. (2012). "Big data": big gaps of knowledge in the field of internet science. *Int. J. Internet Sci.* 7, 1–5.

Smith, K., and Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Hum. Factors* 37, 137–148.

Sonnenwald, D. H., and Pierce, L. G. (2000). Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. *Inform. Process. Manage.* 36, 461–479. doi: 10.1016/S0306-4573(99)00039-4

Strogatz, S. H. (2001). Exploring complex networks. *Nature* 410, 268–276. doi: 10.1038/35065725

Swan, J. M., and Hennig, J. (2012). *From Data to Decisions.* Army Acquisition, Logistics, and Technology (AL and T), 8–12.

Tang, Y., Lebiere, C., Sycara, K., Morrison, D., Lewis, M., and Smart, P. (2015). "Information sharing for collective sensemaking," in *Proceedings of the Hawaii International Conference on System Sciences* (HICSS-49) (Koloa, HI: IEEE Digital Library).

Taylor, R. M. (1990). "Situation awareness rating technique (SART): the development of a tool for aircrew systems design," in *Proceedings of the AGARD AMP Symposium on Situation Awareness in Aerospace Operations* (Neuilly-sur-Seine: Advisory Group for Aerospace Research and Development), 3-1–3-17.

Warwick, W., Buchler, N., and Marusich, L. (2013). "Network science and human performance modeling," in *Proceedings of the Human Factors and Ergonomic Society* (San Diego, CA).

West, B. J. (2012). *Complex Worlds: Uncertain Unequal and Unfair.* Castroville, TX: Black Rose Writing.

White, T. (2012). *Hadoop: The Definitive Guide, 3rd Edn.* Sebastopol, CA: O'Reilly Media, Inc.

Wright, M. C., Taekman, J. M., and Endsley, M. R. (2004). Objective measures of situation awareness in a simulated medical environment. *Qual. Safe. Health Care* 13, i65–i71. doi: 10.1136/qshc.2004.009951

Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K. I., Kondor, I., and Barabási, A. L. (2006). Modeling bursts and heavy tails in human dynamics. *Phys. Rev. E* 73:036127. doi: 10.1103/PhysRevE.73.036127

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* 33, 452–473. doi: 10.1086/jar.33.4.3629752

Zheng, X. S., Tai, Y. C., and McConkie, G. W. (2004). "Exploring drivers' situation awareness in a dynamic traffic environment," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* New Orleans, LA: SAGE Publications.

# APPENDIX

| | Coef. | SE |
|---|---|---|
| **ERGM: Week 1** | | |
| Edges | −4.705 * ** | 0.305 |
| Within-cell homophily | 1.200 * ** | 0.099 |
| Messages sent (in-ties) | −0.001 | 0.000 |
| Messages sent (out-ties) | 0.005 * ** | 0.000 |
| Messages received (in-ties) | 0.004 * ** | 0.001 |
| Messages received (out-ties) | −0.002 * ** | 0.001 |
| Messages sent heterophily | −0.002 * ** | 0.000 |
| Messages received heterophily | −0.002 * ** | 0.000 |
| SA (in-ties) | 0.159 | 0.384 |
| SA (out-ties) | 0.081 | 0.326 |
| SA heterophily | −0.163 | 0.309 |
| Reciprocity | 2.002 * ** | 0.153 |
| Triadic closure (GWESP) | 0.503 * ** | 0.050 |
| GWESP alpha | 1.498 * ** | 0.017 |
| GW outdegree | −0.439 | 0.350 |
| GW indegree | 2.352 * ** | 0.652 |
| | AIC: 3980 | BIC: 4092 |
| **ERGM: Week 2** | | |
| Edges | −5.235 * ** | 0.300 |
| Within-cell homophily | 1.105 * ** | 0.102 |
| Messages sent (in-ties) | −0.004 * ** | 0.001 |
| Messages sent (out-ties) | 0.006 * ** | 0.000 |
| Messages received (in-ties) | 0.007 * ** | 0.001 |
| Messages received (out-ties) | −0.004 * ** | 0.001 |
| Messages sent heterophily | −0.001* | 0.000 |
| Messages received heterophily | −0.003 * ** | 0.001 |
| SA (in-ties) | 1.060 * * | 0.395 |
| SA (out-ties) | −0.989* | 0.341 |
| SA heterophily | −1.351 * ** | 0.331 |
| Reciprocity | 2.080 * ** | 0.149 |
| Triadic closure (GWESP) | 0.750 * ** | 0.061 |
| GWESP alpha | 1.301 * ** | 0.016 |
| GW outdegree | 0.366 | 0.332 |
| GW indegree | 4.535 * ** | 0.848 |
| | AIC: 4110 | BIC: 4223 |

*$* < 0.05, ** < 0.01, *** < 0.001.$*

# Supervising and Controlling Unmanned Systems: A Multi-Phase Study with Subject Matter Experts

*Talya Porat[1,2]\*, Tal Oron-Gilad[2], Michal Rottem-Hovev[3] and Jacob Silbiger[4]*

[1] Department of Primary Care & Public Health Sciences, King's College London, London, UK, [2] Department of Industrial Engineering and Management, Ben Gurion University of the Negev, Beer Sheva, Israel, [3] HFE Independent Consultant, Tel-Aviv, Israel, [4] Synergy Integration Ltd., Tel-Aviv, Israel

Proliferation in the use of Unmanned Aerial Systems (UASs) in civil and military operations has presented a multitude of human factors challenges; from how to bridge the gap between demand and availability of trained operators, to how to organize and present data in meaningful ways. Utilizing the Design Research Methodology (DRM), a series of closely related studies with subject matter experts (SMEs) demonstrate how the focus of research gradually shifted from "how many systems can a single operator control" to "how to distribute missions among operators and systems in an efficient way". The first set of studies aimed to explore the modal number, i.e., how many systems can a single operator supervise and control. It was found that an experienced operator can supervise up to 15 UASs efficiently using moderate levels of automation, and control (mission and payload management) up to three systems. Once this limit was reached, a single operator's performance was compared to a team controlling the same number of systems. In general, teams led to better performances. Hence, shifting design efforts toward developing tools that support teamwork environments of multiple operators with multiple UASs (MOMU). In MOMU settings, when the tasks are similar or when areas of interest overlap, one operator seems to have an advantage over a team who needs to collaborate and coordinate. However, in all other cases, a team was advantageous over a single operator. Other findings and implications, as well as future directions for research are discussed.

Keywords: unmanned aerial systems, control ratio, UAV, decision support systems, DSS, automation, macrocognition, human factors

## INTRODUCTION

The continuing proliferation in the use of UASs in both civil and military operations has presented a multitude of human factors challenges, including assessing the cognitive capabilities of one operator to simultaneously supervise and control multiple platforms, evaluating the advantages and disadvantages of an individual operator vs. a team, and finding meaningful ways to organize and present data. Underlying many of these challenges is the issue of how automation capabilities can best be utilized to assist human operators in handling increasing complexity and workload (Fern et al., 2011).

When the first unmanned aerial systems (UASs) were introduced in the 1980s, engineers and military leaders were content with their ability to extend capabilities of intelligence perception

beyond the capacities that were there before. Once these technological advancements became part of a routine, it became evident that the ratio of personnel vs. crafts issue will rise. There are multiple reasons why managers and leaders are interested in reducing the man-machine control ratio, only to mention a few: fewer operators mean less need for training, less diversity in training, and reduced costs of manpower and training.

The focus on operator-UAS ratio corroborated even more in light of the US Office of the Secretary Defense Roadmap for unmanned aircraft systems (UASs: 2005-2030)[1], which delineates the need to investigate the "appropriate conditions and requirements under which a single pilot would be allowed to control multiple airborne UA [unmanned aircraft] simultaneously." Since then, till today the question of how many UASs or UAVs (Unmanned Aerial Vehicles) can one operator control or supervise has become a vital question that many researchers try to answer (e.g., Chen et al., 2013; Goodrich and Cummings, 2014).

Cummings et al. (2007a) proposed a hierarchical control model to portray control loops for a single operator in control of one UAV or multiple systems. In this three-level model, the innermost loop (Flight controls) represents the need for basic guidance and motion control (i.e., keeping the aircraft in stable flight) and is the most critical. If operators must interact in this loop, the cost will be very high since this loop requires significant cognitive resources. The second loop (Navigation) represents the actions that should be executed to meet mission constraints, such as routes to waypoints, time on targets, and avoidance of threat areas. The outermost loop (Mission and payload management) represents the highest levels of control—decisions which require knowledge-based reasoning that must be made to meet overall mission requirements. Health and status monitoring are tasks that cross all three loops, where the operator is required to perform continuous supervision to ensure that all systems are operating within normal limits. Hence, in order for one operator to be able to control multiple systems, operators will need to interact primarily at the outermost loop via a mission and payload manager while relegating routine navigation and motion control tasks to the automation. For example, given such significant autonomy, one operator could control 4–5 vehicles (Cummings et al., 2007a) and apply supervisory control for up to 12 vehicles (Cummings and Guerlain, 2007).

Higher levels of automation will enable operators to increase the number of unmanned systems they control and supervise, however, extensive use of automation can also introduce human performance costs such as loss of situation awareness, skill degradation, complacency, increased mental workload (Parasuraman et al., 2000) and automation bias (Mosier and Skitka, 1996). Hence, supervisory control of multiple UASs raises questions concerning how to balance system autonomy and human interaction (Calhoun et al., 2011, 2013). Furthermore, the challenge of incorporating automation in one vehicle is replaced by the need to keep the human "in the loop" of the activities for all vehicles (Ruff et al., 2002). Careful system design can

mitigate performance costs and can be achieved by: allowing flexibility in the design of function allocation (i.e., which tasks will be performed by the human and which will be performed by the system), the level of automation to be implemented within each function (Parasuraman et al., 2000; Chen et al., 2013; Gu et al., 2014), and the operators' level of trust in the automation (Clare et al., 2015). Eventually, when flight control becomes fully automated, operators will manipulate the payloads rather than fly the vehicles (e.g., Cooper and Goodrich, 2008).

Ruff et al. (2002) compared the effects of automation level and decision-aid fidelity on the number of simulated remotely operated vehicles (ROVs) that could be successfully controlled by a single operator during a target acquisition task. Their results indicated that an automation level incorporating management-by-consent had clear performance advantages over the more autonomous (management-by-exception) and less autonomous (manual control) levels of automation. Calhoun et al. (2011) used a UAV simulation environment to evaluate two applications of autonomy levels across two primary control tasks: allocation (assignment of sensor tasks to vehicles) and router (determining vehicles' flight plans). Their results showed that performance on both primary tasks and many secondary tasks was better when the level of automation was the same across the two sequential primary tasks. Thus, having the level of automation similar across closely coupled tasks reduced mode awareness problems, which can negate the intended benefits of a fine-grained application of automation.

Adaptive automation (AA) alters the level of automation dynamically during operation. This allows the automation to account for individual differences and allows the automation to be more flexible, context-dependent, and user-specific (Saqer et al., 2011). Wilson and Russell (2007) demonstrated that the customization of automation and difficulty level to the individual operator had greater potential benefit than AA developed based on group performance means. Cummings et al. (2010) examined the impact of increasing automation re-planning rates on operator performance and workload when supervising a decentralized network of heterogeneous unmanned vehicles. They claimed that the future of one operator controlling multiple UVs requires automated planners, which are faster than humans at path planning and resource allocation. They examined three increasing levels of re-planning, and showed that rapid re-planning can cause high operator workload, ultimately resulting in poorer overall system performance. Calhoun et al. (2013) designed an interface enabling pilots to flexibly change the role of automation during the mission, transitioning between four control modes ranging from manual to high level "plays." Their results showed that this approach is promising for single operator supervisory control of multiple UASs, however participants claimed that flexibility should be increased even more, enabling the operator to employ multiple control modes in a single task.

While automation can definitely increase the number of UASs a single operator can supervise and control, Hancock et al. (2007) raised a concern with the ongoing debate over how many UASs should or can a single operator control. The functional design questions that were raised were: (a) should researchers and designers continue to strive for a higher ratio, and, (b) if they

---

[1]*Office of the Secretary of Defense, "Unmanned Aircraft Systems (UAS) Roadmap, 2005-2030."* Washington DC: DoD.

decide to go forward in this direction, what is the modal number? As with all design questions, the immediate answer was simple: It depends. To be sure, the human being as the ultimate adaptive system may be able to demonstrate multiple UAS control, but we consider this an instance of what design can do, not what design should do. In response, John Senders commented that "*with appropriate control and display systems, the handling of more than one machine remains both useful and practical. Simultaneous (actually, appropriately sampled) control of many high-order systems by one operator was demonstrated to be feasible when the displays of attitude are appropriately quickened. Henry P. Birmingham demonstrated this many decades ago by showing excellent simultaneous control of 2 two-dimensional, third-order systems* (Birmingham and Taylor, 1954[2]) .... *Even modestly intelligent design would allow multiple UAVs and multiple displays to be searched or monitored efficiently with good connectivity between the displays. The individual operator is therefore the appropriate unit of analysis only when such bottlenecks occur at that level. More generally, if one views the collective team as an integrated, flexible system, then the very question of the UAV:Operator ratio itself becomes irrelevant.*"

After decades of field practice, the importance of operational use of UASs in combat and in civil operations has increased tremendously. Different team configurations consisting of Multiple Operators and Multiple UASs (MOMU) are nowadays evaluated (e.g., Mekdeci and Cummings, 2009; Gao et al., 2014), implying that indeed the operator to UAS ratio has become an outcome but not a target of its own.

MOMU is a relatively new operational setup for covering areas of interest, particularly in reconnaissance missions. It is highly relevant for homeland security and surveillance operations. A mode of one operator controlling multiple UASs can often increase the cognitive burden of its operators. MOMU setups aim to prevent high operator workload and low situation awareness, and can be very advantageous in offloading tasks to distribute workload among operators. Furthermore, MOMU setups can be advantageous also in terms of utilization of assets, as they contribute to increasing payload efficiency and system effectiveness. However, MOMU settings initiate new challenges for operators as they require switching of information sources, i.e., tasks, missions, video feeds, or camera manipulations and responsibilities among operators.

Switching is a time-critical, cognitively demanding task. Cognitive costs of switching may be loss of orientation and situation awareness (SA), increase in workload, and decrease in efficiency of verbal team communication and coordination. Consequently, switching between sources can disrupt operator performance (Draper et al., 2008; Squire and Parasuraman, 2010), and generate slower and less accurate responses compared to performing a single type of task (Allport et al., 1994; Monsell,

---

[2]In 1954, Taylor and Birmingham, published a paper in the Journal of the Institute of Radio Engineers (now IEEE), titled "A Design Philosophy for Man-Machine Control Systems" (Birmingham and Taylor, 1954). The article discussed the manual control of a submarine, which is a complex control problem because of the massiveness of the boat and the nature of the control surfaces. They also described "quickening", a clever example of how one could augment the display of information to improve the stability of control.

2003). In MOMU environments, where operators need to handoff aircrafts, payloads, targets, or missions to each other, switches may have a vital effect on mission accomplishment.

Over the past decades our team has advanced and improved operational concepts for UASs operators in surveillance and recon missions. Like most others, our studies began with examining the UAS to operator ratio, then to how to increase capacity of a single operator by utilizing tools and automation modes, which gradually shifted toward the MOMU framework. Here we report and revisit these multi-phase studies. Our goal is to demonstrate how the focus of research and practice moved toward a more collaborative operational concept that enables distribution of work and assets among multiple operators. We demonstrate the progress that has been occurring in this human-unmanned system research and how we perceive it should be further directed. We begin with operator to UAS ratio studies. Then, we demonstrate how the MOMU concept evolved. Lastly, we discuss why the changes in UASs control concepts are relevant for other less mature human-robot control domains.
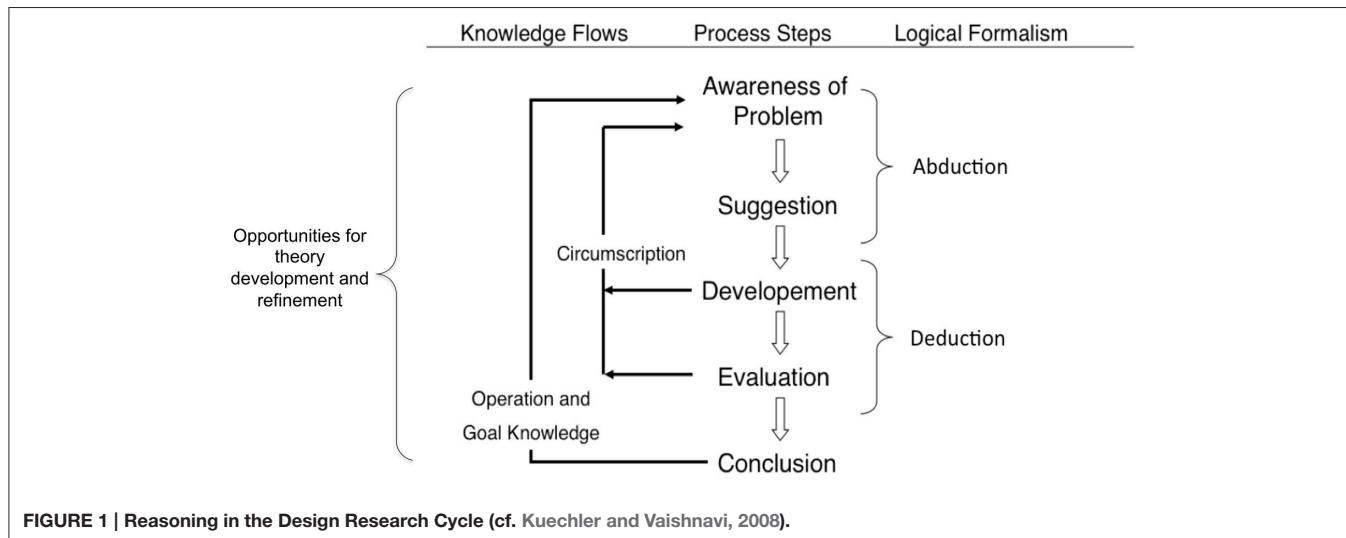
The series of studies has been utilizing the Design Research Methodology (DRM; Blessing and Chakrabarti, 2009). DRM is sometimes called "Improvement Research" emphasizing the problem solving/performance-improving nature of the activity. It enables researchers and analysts to rapidly develop and test prospective improvements, deploy what they have learned about what works, and add to their knowledge to continuously improve the performance of the system (Vaishnavi and Kuechler, 2004). Our aim was to look at the problem from different levels of activity (e.g., supervise, control, mission management), settings (individual vs. team), resources (number of operators, number of vehicles), and automation levels.

In this paper, we do not portray details of every single step in each individual study. Our focus was on the design implications that stemmed from each study phase. This was a conscious strategy, not to be reductionists *per se*, but to allow examination of the operational concept issues from a higher perspective. All the evaluations that are presented were conducted with highly experienced UAS operators (subject matter experts; SMEs) which is necessary for DRM.

## METHODS

We utilized the DRM, with SMEs which focuses on what works, for whom, and under what conditions. In this model (see **Figure 1**) all designs begin with *Awareness of a problem*; then usually from the existing knowledge of the problem area, solutions are *suggested*, after the suggestion phase, there is an attempt to implement an artifact according to the suggested solution—the *Development* phase. Partially or fully successful implementations are then *evaluated* with potential users. *Development*, *Evaluation* and further *Suggestions* are frequently iteratively performed in the course of the research (design) effort. The basis of the iteration, the flow from partial completion of the cycle back to *Awareness of the Problem*, is indicated by the *Circumscription* arrow. *Conclusion* indicates termination of a specific design project. New knowledge production is indicated by the arrows labeled *Circumscription* and *Operation*

**FIGURE 1 | Reasoning in the Design Research Cycle (cf. Kuechler and Vaishnavi, 2008).**

*and Goal Knowledge* (Vaishnavi and Kuechler, 2004; Kuechler and Vaishnavi, 2008). The goal of DRM is to help design research become effective and efficient by making the most out of valuable resources and applying gathered knowledge "on the move." It is particularly suitable for complex interactive systems.

The studies took place at a designated laboratory at Synergy Integration Ltd. which was set up to resemble a typical UAS control room (see **Figure 2**). The work environment was simulated, but "true to life," mimicking UAS military operators' work, who need to operate UASs while placed in a remote designated cabin. The lab consisted of several connected workstations containing a simulation system, which could be configured according to the task and needs (i.e., number of vehicles, individual vs. team operation, time limitations, use of decision support tools, etc.). In this setting, cognitive tasks such as planning, detecting problems, and managing uncertainty (macro-cognitive processes) could be evaluated. Level of automation and mission components were chosen using arrangements similar to the control loops of Cummings et al. (2007a).

## STUDIES

In the following we describe four studies, with their sub-conditions. The earlier two studies examined the operator/platform ratio in several operational scenarios and tasks. The first study examined the number of UASs one operator can <u>supervise</u> (health and status monitoring). The second study examined the number of UASs one operator can <u>control</u> (Mission and payload management) at a single instance. Studies 3 and 4 compared performance of one operator vs. a team of operators controlling the same number of UVs (MOMU studies). Study 3 took place in the UAS environment, while Study 4 took place in the UGV (unmanned ground vehicle) environment. This enabled us to further examine commonalities between the domains of operation. In the following, each study with its different experimental conditions is described.



**FIGURE 2 | The simulated environment.** In the configuration shown here three operators are collaboratively operating three UASs at the same time.
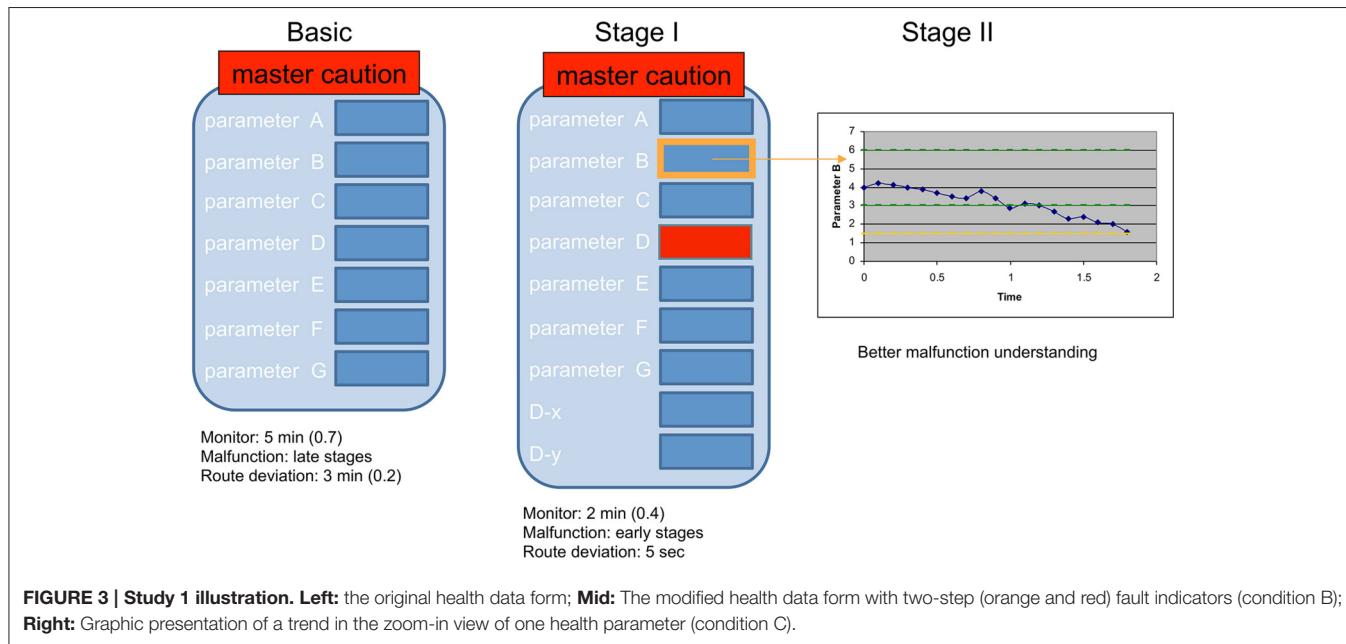
## Study 1

*Problem:* Starting the project, in what may now seem archaic for the UAS domain, health monitoring was identified as the main attention pitfall for operators. Back then, operators had to check the system's health repeatedly while they were performing the flight mission. Displayed health data had to be compared manually against a manufacturer checklist, an error prone process with heavy reliance on memory and specifically prospective memory (see **Figure 3**).

The first study aimed to facilitate the health monitoring task, using automation and tools in order to increase the efficiency and the number of UAV's that one operator could supervise simultaneously.

*Study question:* How many UAVs can one operator supervise (health monitoring) efficiently?

*Participants:* Five highly experienced male operators. All are reserve soldiers in active duty. They had 4–7 years of experience

**FIGURE 3 | Study 1 illustration. Left:** the original health data form; **Mid:** The modified health data form with two-step (orange and red) fault indicators (condition B); **Right:** Graphic presentation of a trend in the zoom-in view of one health parameter (condition C).

in operating military UASs (mean: 5.2), and their age ranged from 23 to 30 (mean: 26.6). SMEs were compensated for their time. The same five participants performed each one of the study conditions, hence a within-subject design was used. Since in DRM one makes incremental design changes, and this process takes time, there was a significant time gap between the different conditions (at least 1 month).

### Initial State—Manual, Sequential Supervising
*Task*
1:5—one operator manually supervised five UAVs of the same kind (utilizing a paper-based checklist).

*Procedure*
For each UAV, 13 health indices were displayed numerically on a form. In addition, two location indices were displayed on a map (X-Y coordinates, related to the pre-defined route). To evaluate the health status of the UAV, participants had to compare the values on the on-line form to a paper-based checklist with the appropriate value ranges. On the screen, the operator could view the health data of only one UAV at a time (i.e., the task required sequential browsing of the health forms). Operators performed continuous manual health monitoring by comparing each index in each form to the desired values written in the hard-copy. While doing this, operators had to relate to different flight stages, as health values varied as a function of flight stage.

*Results*
The cycle time to supervise one UAV was very long—5 min ($SD = 0.7$). The time to detect a fault depended on its location on the form and most faults were detected in late stages of the flaw. Detecting the fault source was almost impossible and took on average 13 min ($SD = 6$). Deviations from the planned route

were detected late, after an average of 3 min ($SD = 0.2$), hence, only after there was a meaningful deviation from the route on the map (scale: 1:50,000).

Operators indicated that the task was difficult and exhaustive within less than 1 h of supervising. They complained on high workload and that they could not imagine succeeding in supervising another (6th) UAV.

### Condition A—Simultaneous Supervising
*Task*
1:5—one operator manually supervised five UAVs with two changes relative to the initial state condition.

*Suggestion—design change from initial state*
To facilitate manual health monitoring, two design implementations were introduced: (1) for each data item an intact indication was added, depending on the flight stage: Intact, Warning (5% lower or higher than the intact value), or Fault; (2) all UAV health data forms were displayed simultaneously.

*Results*
The cycle time to supervise one UAV has decreased from 5 to 2 min ($SD = 0.4$). Most faults were detected in early stages (an average of 5 s to detect a fault). Detection of fault source and route deviations did not improve or differ from the initial state.

### Condition A+—Like A but with More Systems
*Task*
1:10—one operator supervised manually 10 UAVs with the same design as in Condition A.

*Suggestion—design change from Condition A*
Five additional UAVs were added to the supervising task. The limitation to 10 was due to screen size (which enabled displaying up to 10 UAV health forms simultaneously).

## Results

Similar results to condition A—the cycle time to supervise one UAV remained 2 min on average ($SD = 0.4$). Most faults were detected in early stages (average of 5 s to detect a fault). Detection of fault source and route deviations did not improve from the initial state.

## Condition B—Grouping the Health Indices
### Task

1:10–1:20—Operators started with supervising 10 UAVs. During the evaluation, UAVs were added gradually until a single operator was supervising 20 UAVs at a time. To facilitate supervising, the 12 health indices were grouped into four categories.

### Suggestion—design change from Condition A+

There was a change in the display design: the two location indices and one health index were removed (the focus was now only on health parameters). The remaining 12 health indices were grouped into four meaningful categories (e.g., engine, communication, etc.). For each of them, three intact indications were displayed: Intact, Warning and Fault. The shape of the indication icon implied on the contained data in each category. For example, the group containing communication measures (increase/decrease) had an indication icon of arrows pointing up or down.

For each UAV only group indications were displayed on the health data form. The operator could open the full form by clicking on the indication group.

### Results

Results were similar to the ones in condition A. Operators reported upon high workload and a feeling of losing control once the 17th UAV was added.

## Condition B+—Single Indicator for Each System
### Task

1:10–1:20—the operator started supervising 10 UAVs. During the study UAVs were added gradually until stopped at one operator supervising 20 UAVs with a change in the way intact indications were displayed.

### Suggestion—design change from Condition B

The four group indications used in Condition B for each UAV were replaced with one intact indication (icon) for each UAV placed on the command and control map. The operator could click on the icon and view the detailed form. In addition, an alert was added for location deviation.

### Results

Results were similar as in condition A, except for the time to detect deviations from route which was dramatically shortened to 5 s on average (instead of 3 min in previous conditions). Operators succeeded in supervising 15-17 UAVs.

## Condition C—Addition of Malfunction/Health Problem Trends
### Task

1:10–1:20—Operators started with supervising 10 UAVs. During the study UAVs were added gradually until stopped at one operator supervising 20 UAVs. The major change was the addition of a graph display to identify trends in health measures.

### Suggestion—design change from Condition B+

For each indicator, a graph displaying its measured values and intact indications was added. The graph was displayed once the user clicked on the measure value from the health data form. The purpose of this condition was to evaluate if time based information on any specific indication could decrease the time it took for operators to detect the fault source (i.e., aimed to facilitate better malfunction source detection, see **Figure 3**).

### Results

Results were the same as in condition A, except for the major improvement in the time to detect the fault source, which decreased to less than 5 min in 95% of the cases (instead of an average of 13 min in all previous conditions). The ability to view the behavior of the health-related measure over time has helped the operators to understand and detect the source of the fault. The downside of this measure is that it is only suitable for mature systems where the number of faults is relative small, and there is a clear well established link between the health-related measure and its source. Operators succeeded in supervising up to 10 UAVs, mainly because here, more attention was allocated to detecting the source of the fault than previously, and there was not enough time for all the faults to be further examined.

## Study 1 Summary

After performing the first study with its three main conditions, it is possible to claim that one experienced operator can supervise up to 15 UAVs efficiently using the level of automation, the indication tools and the task characteristics described in conditions B and B+. Nevertheless, since health monitoring is only part of mission demands, it was necessary to further investigate the issue of mission and payload management control in Study 2.

# Study 2

*Problem:* The "classical" ratio concern; there was a requirement to increase the number of UAVs that one operator can control.

*Study question:* How many UAVs can one operator control (mission and payload management) efficiently and how can this ratio be improved.

*Participants:* Ten highly experienced male operators (SMEs) with similar military background and skills. They had 3–10 years of experience (mean: 5.6) —7 SMEs in operating military UASs and 3 SMEs in operating other types of military electro-optical sensors. Their age ranged from 23 to 30 (mean: 26). SMEs were compensated for their time.
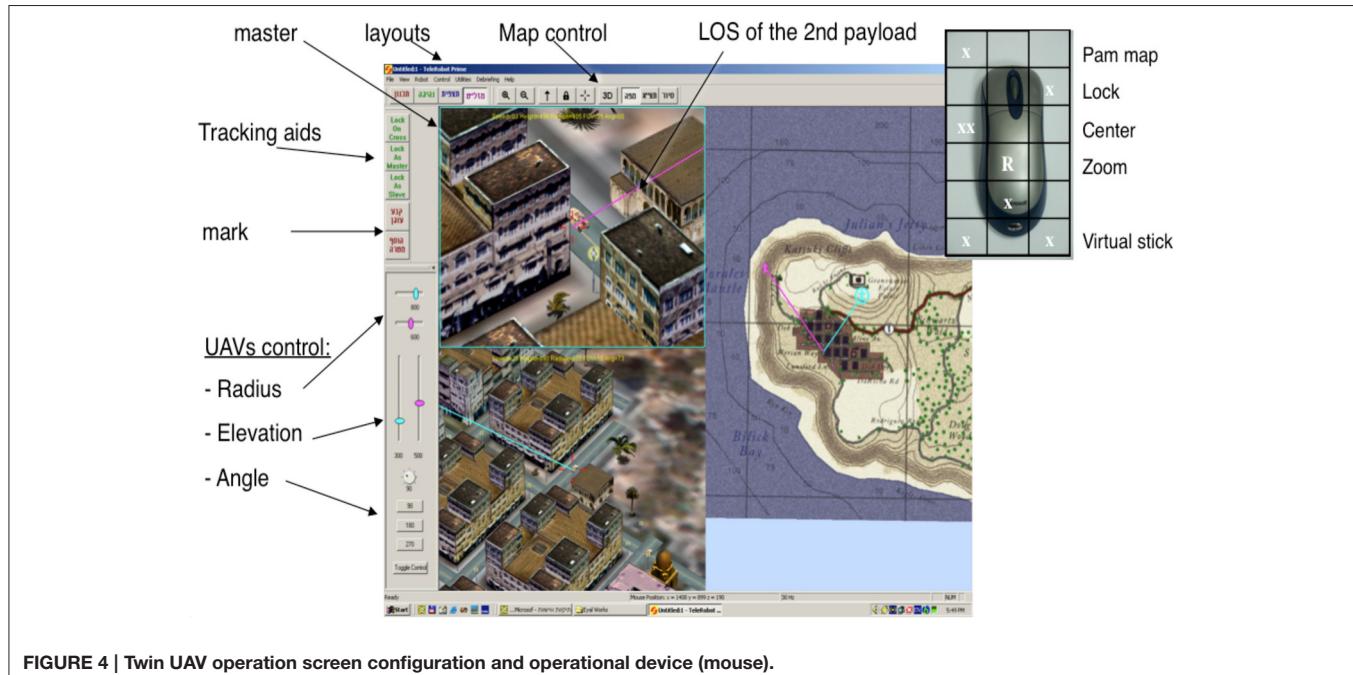
**FIGURE 4 | Twin UAV operation screen configuration and operational device (mouse).**

## Condition A—One to One vs. One to Two

### Task

1:1 vs. 1:2—One operator tracks a moving target with one UAV vs. two UAVs.

Comparing performance of tracking a moving target with two UAVs (Twin UAV setup) vs. a single UAV, in an urban environment. Twin UAV is a "pair of UAVs" handled and operated as one system by one operator (see **Figure 4**). Either UAV can serve as the master while the other one is slaved and vice versa. Hence, only one payload needs to be controlled at a time, and the enslaved UAV positions itself relative to the master. The UAVs control is at a high level of automation via payload management. Various parameters need to be set by the operator for each UAV, prior to each sortie and can be changed during the sortie (altitude, turn radius, camera field-of-view and position shift angle between the UAVs).

### Procedure

The experiment consisted of six experimental scenarios. Each scenario was performed twice, once with one UAV and once with the Twin UAV configuration. The order was counterbalanced among participants. Each trial began with the target vehicle in a specified position. The vehicle then started moving and the operator was asked to keep it in sight as continuously as possible (a lock-on Target feature could be used when the target was visible). Task difficulty depended on the number of similar vehicles in the scene (varied from 5 to 9) and on obstructions when buildings occluded the target. The target vehicle looked similar to other vehicles but had a unique mark. The four easier scenarios lasted 3 min each and the two more difficult ones lasted 4 min. Instructions about the user interface and the task, a demonstration, and four Twin UAV and one single UAV training trials preceded the experimental phase.
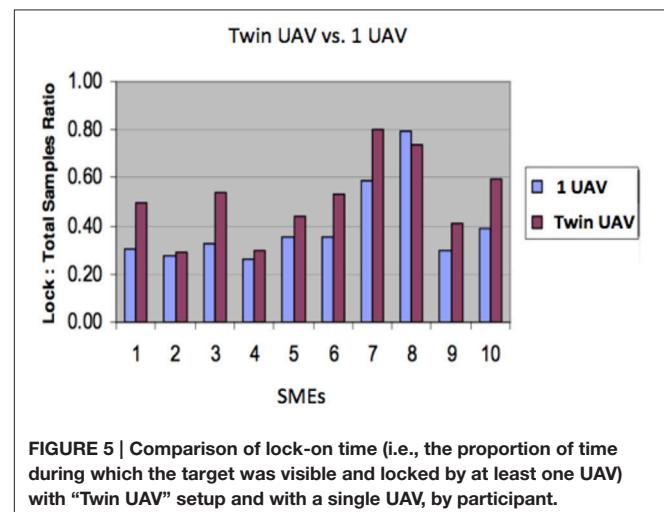


**FIGURE 5 | Comparison of lock-on time (i.e., the proportion of time during which the target was visible and locked by at least one UAV) with "Twin UAV" setup and with a single UAV, by participant.**

### Results

Sampling ratio (time spent in "Lock-on target" mode relative to the total duration of the scenario) was significantly ($p < 0.05$) higher when participants used the Twin UAV (average 0.42, SD 0.12) than the single UAV (average 0.31, SD 0.04). No significant interaction was found between scenario and UAV setup (twin vs. single). **Figure 5** shows the results for each participant.

## Condition B —One to Three

### Task

1:3—Here a more complex operational mission was used; one operator was required to guard a building, track a suspicious vehicle, and scan the shoreline using three UAVs (Tri-UAV), see **Figure 6**.

**FIGURE 6 | Tri-UAV display, the screen is divided into four areas: video feeds of the three UAVs marked with a colored frame for identification (upper left, lower left, and lower right windows), and a command and control map (upper right window).** Note that all three UAVs are shown on the map.

### Procedure

The Tri-UAV display contained video feed windows for each payload and a common map. The operator controlled the display using a mouse and a keyboard. The mouse enabled the operator to move the cursor between the map window and the video feed windows, and point to a specific location.

The task took place in a densely built urban environment. The operator had to: (a) guard a building with several entrances, (b) track a suspicious vehicle, and (c) scan the shoreline. All entrances and exits from and to the building were to be reported. When a suspect vehicle exited the building, the operator had to track it. Two UAVs were allocated to supervising the building entrances while one UAV was used for surveillance (lock-on target to track moving targets could be used). Each scenario was 4 min long and contained eight events that the participant had to attend to, events did not appear at the same time in the scene.

### Results

Operators demonstrated difficulties in simultaneously processing information from three separate locations/video feed sources and failed to succeed in guarding the building and performing additional tasks such as tracking the moving vehicle or scanning the beach line. Only three operators out of the 10 were able to complete the scenarios at some degree of success the remaining seven had difficulties in performing the task and quit before the scenarios ended.

### Study 2 Summary

Experienced operators seemed to cope well with two video feed windows when using the Twin UAV setup. Interestingly, without being instructed to do so, operators intuitively enhanced their performance by utilizing the dual setup. One method that was used frequently by the operators was to choose a wide field-of-view (FOV) angle in one UAV for overview, and a narrow angle on the other UAV for recognition and tracking of the

target. Furthermore, in this type of configuration, since the area of operation was limited, operators rarely used the map. In general, operators thought that handling two sources was difficult enough and that handling three devices may be too demanding. This proved itself correct in condition B, when operators had difficulties processing the information from the three video feed sources. Note also that the area of operation in condition B was wider. In order to succeed, operators stated that there was a need for automated supporting tools. Following these results, in study 3 an attempt was made to facilitate the task by providing the operators with a toolkit containing situation awareness enhancing indicators and decision-support tools.

**Table 1** summarizes studies 1 and 2 as described above. For each study, cognitive task demand, and automation level was added in a separate column (in line with Cummings et al., 2007a). See **Table 2** for the levels of automation legend.

In the following studies performance of a team vs. a single operator was compared in an attempt to understand the feasibility and advantage of each mode, in the UAS domain (Study 3) and in the UGV (unmanned ground vehicle) domain (Study 4). Utilizing the DRM, and based on the findings of the previous studies, tools and visual aids were added to the interface, as specified in each study.

## Study 3

*Problem:* Identify advantages and disadvantages of an individual operator vs. a team. Performance of one operator was compared to a team of (2–4) operators controlling the same number of UAVs (up to four UAVs). Operators had to observe a building and report of vehicles entering and existing the building. Vehicles exiting the building that had specific characteristics had to be further processed.

Study Question: Will a team of operators controlling a number of UAVs perform better than one operator controlling the same number of UAVs?

### Condition A—Two Operators vs. One

*Task*

2:2: vs. 1:2—Two operators sharing control of two UAVs compared to one operator controlling two UAVs.

### Participants

Six highly experienced male operators (SMEs) with similar military background and skills participated in this condition. They had 2–7 years of experience in operating military UASs (mean: 4), and their age ranged from 23 to 27 (mean: 24.8).

### Procedure

Operators had to observe a building and report of vehicles entering and exiting the building. Vehicles exiting the building that had specific characteristics (i.e., suspicious vehicle) had to be further processed (track and report). Two phases were conducted, in the first phase no additional unique interaction tools were provided. After the first phase, based on the findings from study 2 and the difficulties operators had in performing the task, supportive tools were provided, only to the single operator in a form of a toolkit. The toolkit consisted of spatial anchoring

**TABLE 1 | Summary of studies 1 and 2.**

| Study/Condition | Description/Design | Operator/UV ratio | Cognitive task and automation level* | Results |
|---|---|---|---|---|
| **1** | **UAVs HEALTH AND STATUS SUPERVISING** | | | |
| Initial state | Supervise 13 health indices using a paper-based checklist. | 1:5 | Health monitoring-I | Operators indicated that the task was difficult, exhaustive, and caused high workload. |
| A | Two design additions:<br>1. For each data item an intact indication was added<br>2. All UAV forms were displayed simultaneously. | 1:5 | Health monitoring -III | Improved performance from the Initial State. |
| A+ | Display was the same as in condition A. However, five additional UAVs were added to the monitoring task. | 1:10 | Health monitoring -III | Similar results to condition A. |
| B | 12 health indices were grouped into four meaningful groups. For each UAV, only the group indications were displayed on the form. For each group, three intact indications were displayed (intact, warning, and fault). The operator could open the full form by clicking on the indication group. | 1:20 | Health monitoring -III | Similar results to condition A. Operators reported high workload and a feeling of losing control after the 17th UAV was added. |
| B+ | The group indications used in Condition B for each UAV were replaced with one intact indication (icon) placed on the command and control map. The operator could click on the icon and view the details. In addition, an alert was added for location deviation. | 1:20 | Health monitoring -III | Similar results to condition A, except for the time to detect deviation from route which was shortened. Operators succeeded in supervising 15–17 UAVs. |
| C | For each indicator, a graph displaying its measured values and intact indications was added. | 1:20 | Health monitoring-III | Similar results to condition A, except for the time to detect the fault source, which was shortened. Operators succeeded in supervising up to 10 UAVs. |
| **2** | **UAVs CONTROL—1-TO-1 AND 1-TO-MANY** | | | |
| A | Tracking a vehicle once with one UAV and once with a Twin UAV. Target lock-on could be used. | 1:1 vs.1:2 | Navigation—V | Performance was significantly better using two UAVs. |
| B | The operator had to guard a building with several entrances, track a suspected vehicle (target lock-on could be used), and scan the beach line. 2 UAVs were required for supervising the building while one UAV was required for surveillance. | 1:3 | Navigation—V Mission Management-II | Operators demonstrated difficulties in processing information from three separate sources. |

*See **Table 2** for the levels of automation legend.

capabilities like "sketch" and "revisit," which enabled the operator to request the system to automatically follow a pattern (perform a sketch) or a jump through a list of points (perform a revisit cycle) by generating (using mouse clicks) a list of points on top of the payload image. In a similar way to Study 2's Twin UAV setup "Payload coupling" enabled the operator to enslave one UAV to the other. Finally, "Camera guide" enabled the operator to fly the UAV by following its camera (See Oron-Gilad et al., 2011 for detailed description of several tools).

In phase 2 of the study, it was aimed to examine whether the toolkit could support the single operator's performance to a degree superior to the team of two operators.

### Results
Results are displayed in **Table 3**.

The team reported that the mission was calm up to a degree of being boring. The single operator reported that the mission was challenging but not overloading. The results of the team were similar to the results of the single operator using a toolkit. Multiple reporting of the same incident and longer mission stabilization time occurred in the team condition.

### Condition B—Three Operators vs. One
*Task*
3:3 vs. 1:3—A team of three operators sharing control over three UAVs were compared to one operator controlling three UAVs. The same scenarios as in Condition A, the individual operator could use the toolkit and the operators in the team could not.

### Participants

Eight highly experienced male operators (SMEs) with similar military background and skills participated in this condition. They had 4–8 years of experience in operating military UASs (mean: 5.4), and their age ranged from 25 to 30 (mean: 26.9). SMEs were compensated for their time.

### Results

Results are displayed in **Table 4**.

The team performed significantly better ($p < 0.01$) than the single operator, however they again had more occasions of multiple reporting of the same incident, and increased stabilization time.

## Condition B+—Four Operators vs. One

### Task

4:4 vs. 1:4—A team of four operators sharing control of four UAVs were compared to one operator controlling four UAVs.

### Participants

Five highly experienced male operators (SMEs) with similar military background and skills participated in this condition. They had 3–5 years of experience in operating military UASs (mean: 3.83), and their age ranged from 25 to 27 (mean: 25.5).

### Results

This setup was problematic to analyze. In the one operator condition, single operators felt lost looking at four video feeds and in some cases they just looked at three UAVs or less (hence they neglected the fourth UAV). In the team condition, coordination among the operators took a long time, containing incessant verbal communication, and numerous multiple reports.

### Study 3 Summary

One operator could not control more than three UASs, even with additional aids. Furthermore, without facilitating decision support tools, it was difficult and ineffective for a team of four operators to control four UASs as well. The implications of this study were twofold: each single operator can benefit from designated tools that assist in conducting the mission, e.g., coupling or sketch and revisit. A team of operators must be familiarized with a set of rules or provided with a set of tools to facilitate collaboration. Otherwise, they are prone to report multiple times on the same incident and they are not fully aware of each other's doings. Following these findings several novel tools and displays were designed to facilitate payload switching among members of the team (see for example Porat et al., 2011). Probably, the most successful facilitating tool was the "Castling Rays," which is a switching decision aid, enabling operators to visually view which UAS has the best view of "their" target at any given moment (Porat et al., 2010).

## Study 4
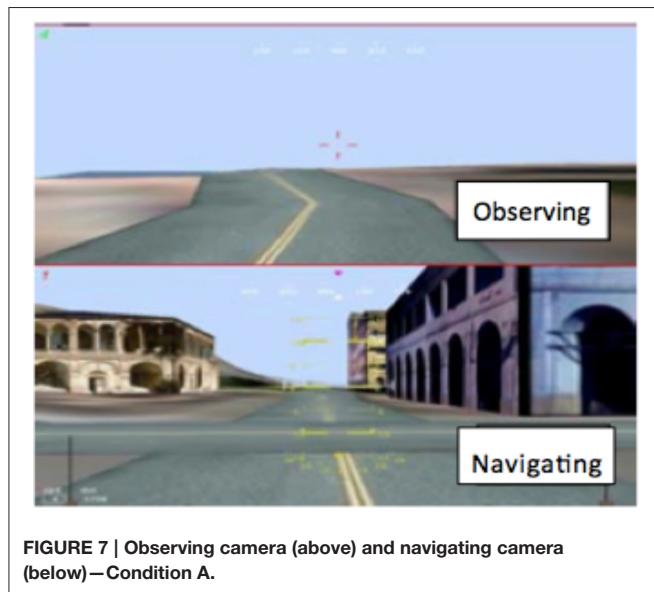
*Problem:* There was a requirement to increase the number of UGVs that one operator can control. The main problem with UGVs is that their level of autonomy is lower, hence more attention needs to be allocated to navigation and driving issues than in UAVs. At the time tested, the problem domain was still

**TABLE 2 | Levels of Automation (LOA) (cf Cummings et al., 2007a).**

| SV*-LOA | C**–LOA | Automation Description |
|---|---|---|
| 1 | I | The computer offers no assistance; human must take all decisions and actions. |
| 2 | II | The computer offers a complete set of decision/action alternatives. |
| 3 | III | The computer offers a selection of decisions/actions. |
| 4/5 | IV | The computer suggests one alternative and executes that suggestion if the human approves (management by consent). |
| 6 | V | The computer suggests one alternative and allows the human a restricted time to veto before automatic execution (management by exception). |
| 7/8/9/10 | VI | The human is not involved in the decision making process; the computer decides and executes autonomously. |

*SV—The 10-level scale originally proposed by Sheridan and Verplank (1978).*
**C—The combined categories of Cummings et al. (2007a).*

**TABLE 4 | Performance measures—Team of 3 vs. one operator controlling 3 UAVs.**

| Performance measure | Team of UAVs 3—3 UAVs | UAVs + toolkit 1 operator—3 |
|---|---|---|
| Misses of vehicle exits | 1% | 3% |
| Misses of vehicle entrances | 1% | 3% |
| Misses of a suspect vehicles | 0% | 0.5% |
| Multiple reporting of the same vehicle | 25% | 8% |
| Mission stabilization time | 13 min (SD 3) | 3.5 min (SD 1.2) |

**TABLE 3 | Performance measures—Team of 2 vs. one operator controlling two UAVs.**

| Performance measure | Team of 2—2 UAVs | 1 operator—2 UAVs | 1 operator—2 UAVs + toolkit |
|---|---|---|---|
| Misses of vehicle exits | 2% | 2% | 3% |
| Misses of vehicle entrances | 3% | 4% | 3% |
| Misses of a suspect vehicles | 0.5% | 0.7% | 0.5% |
| Multiple reporting of the same vehicle | 15% | 5% | 5% |
| Mission stabilization time | 10 min (SD 2.5) | 6 min (SD 1.75) | 2 min (SD 0.8) |

**FIGURE 7 | Observing camera (above) and navigating camera (below)—Condition A.**

**TABLE 5 | Performance measures of the initial state.**

| Performance measure | Initial state |
|---|---|
| Time to identify an obstacle | 4 s (SD 1.5) |
| Missing an obstacle* | 3% |
| Time to identify a hazard on the fence | 3 s (SD 1.5) |
| Missing a hazard on the fence | 2% |

*All missed obstacles were of type "pitfall." Pitfalls are more difficult to identify than above ground hazards such as a log put on the ground.*

**TABLE 6 | Performance measures—initial state vs. condition A.**

| Performance measures | Initial state | Condition A |
|---|---|---|
| Time to identify an obstacle | 4 s (SD 1.5) | 7 s (SD 2.3) |
| Missing an obstacle* | 3% | 6% |
| Time to identify a hazard on the fence | 3 s (SD 1.5) | 9 s (SD 3) |
| Missing a hazard on the fence | 2% | 5% |

*All missed obstacles were of type "pitfall." Pitfalls are more difficult to identify than above ground hazards such as a log put on the ground.*

within the realm of multiple operators controlling a single system vs. a single operator. We compared performance of two operators controlling (navigating and observing) one UGV to one operator controlling one UGV.

*Study Question:* Will two operators controlling one UGV perform better than one operator controlling one UGV for scanning the fence?

## Initial State
### Task
2:1—Two operators controlled one UGV: one operator performed the navigation task and one operator performed the observation task, while scanning a border fence.

### Participants
Six highly experienced participants, reserve soldiers in an elite engineering unit with experience in controlling remote robots such as ANDROS and Mini-ANDROS participated in this condition. They had 2–6 years of experience in operating military UVGs (mean: 3.5), and their age ranged from 25 to 30 (mean: 26.8). All were compensated for their time.

### Procedure
Each UGV had a navigation camera and an observation camera for scanning the fence for obstacles and hazards (**Figure 7**). The UGV moved very slowly (7 km/h). One operator performed the navigation task (including health monitoring—alerts were both color coded and audible), and one performed the observation task. The experimental trial took about an hour. In this period, a total of 100 events occurred (obstacle, hazard on the fence, fault in the vehicle).

### Results
Results are displayed in **Table 5**.

Performance was acceptable with a relatively low rate of misses of obstacles. However, there were synchronization problems among the two operators, for example: operators had delays in stopping the vehicle, which usually occurred after the observer identified a hazard, and notified the navigator who then had to stop the vehicle.

## Condition A
### Task
1:1—one operator controlled one UGV, performing both the navigation and the observation tasks (as shown in the display in **Figure 7**).

### Participants
Three highly experienced participants, reserve soldiers in an elite engineering unit with experience in controlling remote robots such as ANDROS, and Mini-ANDROS participated in this condition. They had 2–4 years of experience in operating military UGVs (mean: 2.7), and their age ranged from 25 to 28 (mean: 26.3).

### Results
Performance measures between the "Initial State" and "Condition A" were compared. Results are displayed in **Table 6**.

One of the main problems in this condition was that operators were missing pitfalls, which stopped the vehicle and increased the time based performance measures to a large extent.

## Condition B
### Task
1:5—one operator **observed** cameras from five different UGVs, scanning the fence for obstacles and hazards.
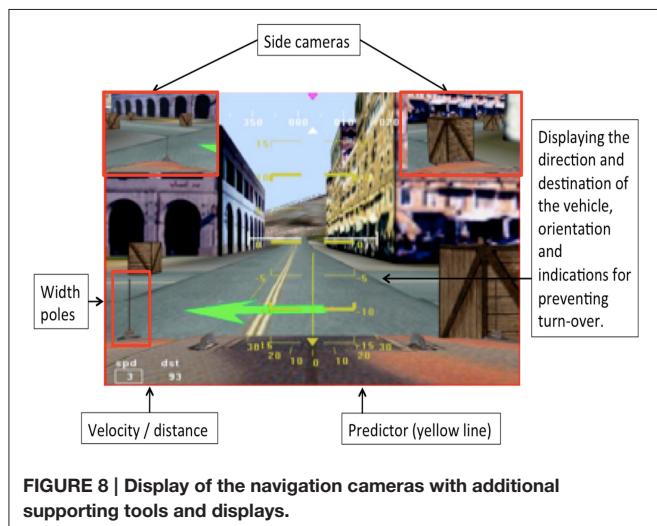
### Participants
The same participants as in condition A.

TABLE 7 | Performance measures—initial state vs. condition B.

| Performance measures | Initial State | Condition B |
|---|---|---|
| Time to identify an obstacle | 4 s (SD 1.5) | 5 s (SD 3) |
| Missing an obstacle* | 3% | 4% |
| Time to identify a hazard on the fence | 3 s (SD 1.5) | 4.5 s (SD 2.5) |
| Missing a hazard on the fence | 2% | 3% |

*All missed obstacles were of type "pitfall." Pitfalls are more difficult to identify than above ground hazards such as a log put on the ground.



FIGURE 8 | Display of the navigation cameras with additional supporting tools and displays.

## Results

Performance measures between the "Initial State" and "Condition B" were compared. Results are displayed in **Table 7**.

## Study 4 Summary

It was too complicated for one operator to perform the observation and navigation tasks simultaneously (as in Condition A). These two task types require different skills and performing them at the same time generated major switching costs. However, when operators were performing only one type of task (observation or navigation), their performance has improved.

Based on these findings, several novel tools and displays were designed to facilitate the navigation task, as shown in **Figure 8**. Side cameras were added. A width pole display aided the operator in estimating the width of the vehicle, and a path Predictor displayed a virtual path that the navigator could follow. Initial examination found this setup to decrease navigation time and improve navigation accuracy. This needs to be further assessed, however could be extremely useful especially when there are communication delays in displaying the online video feed from the navigation cameras.

**Table 8** summarizes studies 3 and 4 as described above. For each study, cognitive task demand and automation level were added in a separate column (in line with Cummings et al., 2007a). See **Table 2** for the levels of automation legend.

## SUMMARY AND DISCUSSION

In general, our results suggest that one experienced operator can supervise (system health and status) up to 15 UASs efficiently using moderate levels of flight control automation. Concerning controlling UASs (mission and payload management), one experienced operator cannot control more than three UASs, with the level of complexity and automation that has been examined. Providing the operator with various display aids and decision support tools does improve performances of a single operator (as in Study 3) but did not change the modal number to higher extents.

Automation level, availability of decision aids, operators' experience, complexity and criticality of the mission, operational tempo, and cognitive resources and demands, all influence the number of systems that one operator can control. For this reason, comparison across studies is often complicated and inaccurate. However, considering these limitations, our findings do resemble findings of previous studies in the essence that they are confirming that single operators are able to control more remote vehicles as they are provided with increasing automated decision support. Given some automated navigation assistance and management-by-consent automation in the mission management loop, an operator was able to control 4–5 vehicles (e.g., Ruff et al., 2004; Dunlap, 2006; Cummings et al., 2007b). A leap in the amount of vehicles that one operator could control was only seen if management-by-exception was introduced, increasing the number to 8–12 vehicles (e.g., Lewis et al., 2006; Cummings and Guerlain, 2007). Here, we were able to show via Study 1 that a single operator can achieve even a higher ratio of operation between 15 and 17 systems, but only on a limited task or mission component (e.g., health monitoring).

This finding may become more relevant in the future, if organizations change the way they allocate and recruit operators. Nowadays, most organizations, military amongst them, do not want to parse their operators' mission into "small" subtasks and create high levels of skills in fine grained subtasks of the mission among operators (i.e., train people to be experts only on a single component of the mission, such as taxi or health supervising). The current approach can be justified when considering the danger of having operators lacking skill while conducting dynamic, time critical, and situation critical missions. However, the way operators' allocations are done today, it is inevitable for operators to maintain a certain level of proficiency in all aspects of their mission. Evidently this setting dictates that the level of automation of the unmanned system and the use of decision aids become key considerations.

Human operators are vital in this critical, high risk and high demand environment. Keeping the human in the loop, mostly for planning, re-planning, and control or at least for being able to take over in automation malfunction is essential in this domain. Therefore, fully autonomous operations (automation level VI) are not expected any time soon. Using intermediate levels of automation (i.e., supervisory control), will not enable operators to exceed the control of few systems.

**TABLE 8 | Summary of studies 3 and 4 (MOMU).**

| Study/Condition | Description/Design | Operator/UV ratio | Cognitive task and automation level* | Results |
|---|---|---|---|---|
| **3** | **UAVs CONTROL—1-TO-MANY VS. MANY-TO-MANY** | | | |
| A | The operator had to observe a building and report of vehicles entering and exiting the building. Vehicles exiting the building that had specific characteristics had to be further processed. The single operator could use a toolkit containing decision-making tools. | 1:2 vs.2:2 | Team: Mission Management-I Single: Mission Management-I + Tools | Teams of two operators described the mission as calm (and even boring). The single operators reported that the mission was challenging but not overloading. |
| B | The same scenarios as in Condition A. The single operator could use a toolkit and the team of operators could not. | 1:3 vs.3:3 | Team: Mission Management-I Single: Mission Management-I + Tools | The team performed better than the single operator. However, the team had more double reporting. |
| B+ | The same scenarios as in Condition A. The single operator could use a toolkit and the team of operators could not. | 1:4 vs.4:4 | Team: Mission Management-I Single: Mission Management-I + Tools | This setup could not be examined since the tasks were too complex. |
| **4** | **UGVs CONTROL—MANY-TO-1 AND 1-TO-1** | | | |
| Initial | One operator performed the navigation task and one performed the observation task. | 2:1 | Navigator: Navigation-I Observer: Mission Management-I | Performance was satisfactory. Synchronization problems between the two operators were apparent. |
| A | One operator performed the navigation and the observation tasks. | 1:1 | Navigation-I Mission Management-I | It was too complicated for one operator to perform both observation and navigation tasks. |
| B | One operator observed five UGVs. | 1:5 | Mission Management-I | Performance was good and similar to the initial condition. |

*See **Table 2** for the levels of automation legend.

**FIGURE 9 | Left—operator capacity as a function of mission constrains (cf Cummings et al., 2007b).** Right—Impact of UAS Number from an OR study conducted in parallel to our studies (Shaferman and Shima, 2009).
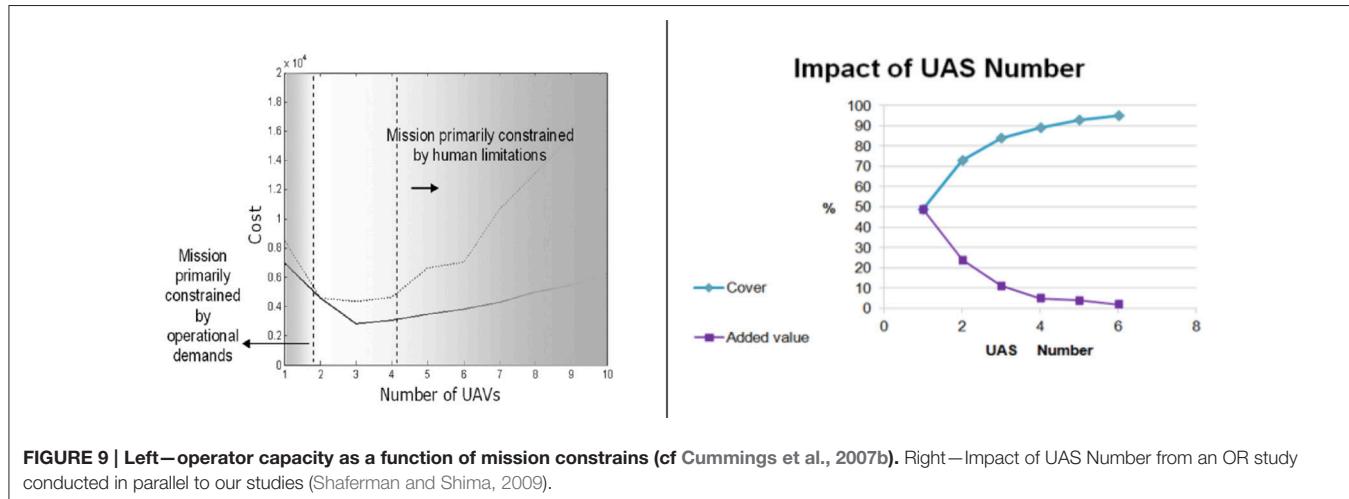
**Figure 9** on the left was taken from Cummings et al. (2007b) and shows that the optimal bound they found was between 2 and 4 vehicles. The left region is primarily constrained by operational demands, but the right region is dominated by human performance limitations. **Figure 9** on the right is taken from an operation research study conducted in parallel to Studies 1–4 and on similar urban area conditions (Shaferman and Shima, 2009). It shows that adding the first and second UAV had the most significant influence on mission performance. Above four systems, the area covered, and the added value of more assets became negligible. Hence, organizations need to identify whether there are justified operational cases where one-to-many ratios of more than four are needed. If those cases are sparse, then perhaps more design effort can be geared toward sharing of assets among operators (MOMU) in an efficient and effective way.

Concerning the operation of UGVs, when the operator performed only one task, as in study 4, condition B (observation task), performance was satisfactory since the operator focused primarily on maintaining awareness for obstacles and hazards. However, when the operator had to navigate the vehicle and observe the fence (as in study 4, condition A), it was too complicated to perform. Dynamic task switching between different functions resulted in greater cognitive workload for the operator than performing only one type of task. In both UASs and UGVs, the human and the automated systems are geographically separated, and therefore face difficulties, which are inherent in remote perception, such as overcoming the "keyhole" or "soda straw effect" (Voshell et al., 2005). Controlling and navigating UGVs is more complex than UASs with regard to spatial perception. While GPS technology may be very effective in providing UASs with positioning information that meets their navigational needs, their use in UGVs may be limited by reliability and accuracy constraints (Chaimowicz et al., 2005). For example, a positioning error of one or two meters may have little effect when controlling a UAS, however it could have crucial results when navigating a UGV.

Successful interaction with any human and automated system is influenced by many factors including vehicle characteristics (air, ground), task characteristics (complexity, number of vehicles controlled, time pressure, workload), environmental characteristics (terrain characteristics, quality, obstacles), and technological constraints (available bandwidth, communication delays). Thus, design specifications of automated decision support aids will differ according to the unique needs of the human operator in each situation. Indeed, the decision support tools that were developed in this study for the aerial and the ground domain differ in their design and implementation (e.g., width pole display for the ground vehicle) but there are also many commonalities in the essence of things (e.g., coupling of vehicles is suitable for both aerial and ground vehicles).

In MOMU environments, as seen in Study 3, when the tasks are similar or when the interest areas overlap (i.e., a connection between the video feeds), one operator has an advantage to a team who need to collaborate and coordinate. However, when there is no connection between the video feeds, a team has an advantage to a single operator. Thus, one of the considerations to prefer one operator to a team is the amount of overlap between the different video sources covered by the payloads. Taking this findings to a practical level, in the MOMU operational settings we strive to gain a consistent ratio of one operator controlling two UASs with some flexibility, thus controlling up to three UASs per operator on demand, and supervise up to six UASs where the covered areas of the UASs are related.

## WHERE CAN WE GO FROM HERE AND BROADEN THE UNDERSTANDING AND ADDED VALUE OF MOMU ENVIRONMENTS?

The first notion is that automation is a tricky tool. When not tailored to the task, it can easily cause high operator workload, and challenge the "keep the human in the loop" principle.

Although this statement may seem true for most human system interfaces, when applying automation in critical and complex environments, such as MOMU, a first step would be to perform a thorough behavioral and cognitive task analysis to understand the cognitive requirements of the task (e.g., decisions, situation awareness, cues, judgment points). Once the different tasks, requirements and possible errors are understood, tailoring the display design and the automation level to the desired setup becomes possible. It should be acknowledged that different sections/parts/sub-tasks of the entire mission are perceived differently at separate stages of the mission process. For example, different automation needs are required for locomotion between areas of interest, as opposed to loitering on a specific target area. This implies that Cummings et al. (2007a) control loops could be further divided into even smaller chunks, and for each chunk one should match the required and desired automation level.

The second notion is that the scenarios used in our studies assumed similarity: all operators had the same type of experience and training, and all systems were alike. While this is a typical mode of operation, it is evident that this is just one possibility. In the U.S. military operations in Iraq, for example, more than 100 UASs of 10 different types were used (Office of the Under Secretary of Defense for Acquisition, 2004). The rising question becomes how MOMU operations may vary if there were multiple types of vehicles and operators with various training and capabilities. One needs also to reconsider the traditional mission allocation. Recent studies tried to define the qualifications and training required from an operator that is expected to control an increasing number of UASs. Parasuraman et al. (2014) discussed the possibility of selecting and training operators according to their molecular genetics. Perhaps now is the time to initiate specialization of operator roles. In order to do so, it would be necessary to revisit the main operational tasks and reallocate them in view of mission benefit. Changes in the function allocation and the nature of task differentiation between human operators and unmanned systems could significantly alter the cognitive loads of the operators when performing the mission (Cuevas et al., 2007). We should introduce flexibility into our rigid-traditional "task thinking," and let go of beliefs that tie us down and stop the evolvement: must human operators fly the platform? Can we mentally not technically—enslave the platform to the mission needs?

A third direction would be to develop tools and decision-making aids. In our studies tools and techniques that may facilitate operators in MOMU environments were introduced (e.g., Porat et al., 2010). Tool development was done in a bottom up approach, i.e., based upon needs retrieved from SMEs and geared toward solving particular challenging operational situations. Since the tools were not yet tested in real world settings, it would be interesting to examine how they integrate into UASs MOMU environments and affect the metrics of

performance. Fern et al. (2011) for example proposed other alternatives to facilitate UAS MOMU operations. It would also be interesting to examine whether tools can be transferred to other MOMU settings such as ground vehicles or drones.

Fourth, our studies focused on the allocation among operators in one team while conducting a single scenario, one can start looking at the broader picture—how to break operations into teams, how to assign and allocate the correct number of assets and operators to each one of the teams, and how to coordinate among teams of MOMU operators.

All these former suggestions lead toward the notion that a more top-down approach needs to be developed in order to provide a coherent way to distribute responsibilities and tasks in MOMU environments. This direction of adjusting resources and personnel according to mission needs is in line with future intentions and models in other domains. For example, in the medical domain, the NHS recent report "Five year forward view" (NHS England, 2014), argues that England is too diverse for a "one size fits all" care model, services need to be integrated around the patient and support their changing needs. Different local health communities will be able to decide which care delivery model best supports their needs, such as a multispecialty community provider model which is a multi-disciplinary team that can include different specialties such as nurses, therapists and other professionals combined with the latest digital technologies, or a specialized care model which is a surgery that specializes in one area such as cancer and provides care only for these patients. All to support the main goal, which is providing the best care for patients. Translating this to our domain and the task specific requirements, we can reach to the extreme cases where a team of operators will control only one asset and vice versa, where a single operator will control up to 15 assets simultaneously (e.g., taxi).

Finally, with regard to Human-Robot Interaction, it is inevitable that people of various abilities and skills will be surrounded by multiple platforms of various kinds and autonomy levels. Much of what is now known from the realm of UASs can be used to facilitate efficient asset sharing and mission successes among other populations. Just to mention one, in the not so far future, the elderly community will be utilizing robotics assistants of various kinds, whether operated by caregivers or by the users themselves. Many of the questions that were raised here about operators' skills, tools to facilitate cooperation and sharing and mission accomplishment will be relevant to these domains as well.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: MR, JS. Analyzed and interpreted the data: TP, TO, MR, JS. Wrote the paper: TP, TO, MR.

# REFERENCES

Allport, D. A., Styles, E. A., and Hsieh, S. (1994). "Shifting intentional set: exploring the dynamic control of tasks," in *Attention and Performance IV,* eds C. Umilta and M. Moscovitch (Cambridge, MA: MIT Press), 421–452.

Birmingham, H. P., and Taylor, F. V. (1954). "A design philosophy for man-machine control systems," in *Proceedings of the Institute of Radio Engineers,* Vol. 42 (IEEE), 1748–1758. doi: 10.1109/jrproc.1954.274775

Blessing, L. T. M., and Chakrabarti, A. (2009). *DRM, A Design Research Methodology.* London: Springer-Verlag. doi: 10.1007/978-1-84882-587-1

Calhoun, G., Draper, M., Miller, C. A., Ruff, H., Breeden, C., and Hamell, J. (2013). "Adaptable automation interface for multi-unmanned aerial systems control: preliminary usability evaluation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 57 (San Diego, CA), 26–30. doi: 10.1177/1541931213571008

Calhoun, G. L., Ruff, H. A., Draper, M. H., and Wright, E. V. (2011). Automation-level transference effects in simulated multiple unmanned aerial vehicle control. *J. Cogn. Eng. Decis. Making* 5, 55–82. doi: 10.1177/1555343411399069

Chaimowicz, L., Cowley, A., Gomez-Ibanez, D., Grocholsky, B., Hsieh, M. A., Hsu, H., et al. (2005). "Deploying air-ground multi-robot teams in urban environments," in *Multi-Robot Systems. From Swarms to Intelligent Automata Vol. III* (Springer), 223–234. doi: 10.1007/1-4020-3389-3_18

Chen, T., Campbell, D. A., Coppin, G., and Gonzalez, F. (2013). "Management of heterogeneous UAVs through a capability framework of UAV's functional autonomy," in *15th Australian International Aerospace Congress (AIAC 15)* (Melbourne, VIC: Melbourne Convention Centre).

Clare, A. S., Cummings, M. L., and Repenning, N. P. (2015). Influencing trust for human-automation collaborative scheduling of multiple unmanned vehicles. *Hum. Factors* 57, 1208–1218. doi: 10.1177/0018720815587803

Cooper, J., and Goodrich, M. A. (2008). "Towards combining UAV and sensor operator roles in UAV-enabled visual search," in *Presented at HRI 08'*, (Amsterdam). doi: 10.1145/1349822.1349868

Cuevas, H. M., Fiore, S. M., Caldwell, B. S., and Strater, L. (2007). Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviat. Space Environ. Med.* 78, B63–B70.

Cummings, M. L., Andrew Clare, A., and Hart, C. (2010). The role of human-automation consensus in multiple unmanned vehicle scheduling. *Hum. Factors* 52, 17–27. doi: 10.1177/0018720810368674

Cummings, M. L., Bruni, S., Mercier, S., and Mitchell, P. J. (2007a). Automation architecture for single operator, multiple UAV command and control. *Int. C2 J.* 1, 1–24.

Cummings, M. L., and Guerlain, S. (2007). Developing operator capacity estimates for supervisory control of autonomous vehicles. *Hum. Factors* 49, 1–15. doi: 10.1518/001872007779598109

Cummings, M. L., Nehme, C. E., Crandall, J., and Mitchell, P. (2007b). "Predicting operator capacity for supervisory control of multiple UAVs," in *Innovations in Intelligent Machines-1,* eds J. S. Chah, L. C. Jain, A. Mizutani, and M. Sato-Ilic (Berlin; Heidelberg: Springer), 11–37. doi: 10.1007/978-3-540-72696-8_2

Draper, M., Calhoun, G., Ruff, H., Mullins, B., Lefebvre, A., Ayala, A., et al. (2008). "Transition display aid for changing camera views in UAV operations," in *Proceedings of the Humans Operating Unmanned Systems*. Available online at: http://conferences.telecom-bretagne.eu/

Dunlap, R. D. (2006). "The evolution of a distributed command and control architecture for semi-autonoimous air vehicle operations," in *Presented at Moving Autonomy Forward Conference* (Grantham).

Fern, L., Shively, J., Draper, M., Cooke, N. J., Oron-Gilad, T., and Miller, C. A. (2011). "Human-automation challenges for the control of unmanned aerial systems," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Vol. 55 (Las Vegas), 424–428. doi: 10.1177/1071181311551087

Gao, F., Cummings, M. L., and Solovey, E. T. (2014). Modeling teamwork in supervisory control of multiple robots. *IEEE Trans. Hum Machine Syst.* 44, 441–453. doi: 10.1109/THMS.2014.2312391

Goodrich, M. A., and Cummings, M. L. (2014). "Human factors perspective on next generation unmanned aerial systems," in *Handbook of Unmanned Aerial Vehicles*, eds K. P. Valavanis and G. J. Vachtsevanos (Springer), 2405–2423.

Gu, W., Mittu, R., Marble, J., Taylor, G., Sibley, C., Coyne, J., et al. (2014). "Towards modeling the behavior of autonomous systems and humans for trusted operations," in *2014 AAAI Spring Symposium Series* (Stanford).

Hancock, P. A., Mouloua, M., Gilson, R., Szalma, J., and Oron-Gilad, T. (2007). Is the UAV control ratio the right question? *Ergonom. Design* 15, 7; 30–31. doi: 10.1177/106480460701500104

Kuechler, B., and Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. *Eur. J. Inf. Syst.* 17, 489–504. doi: 10.1057/ejis.2008.40

Lewis, M., Polvichai, J., Sycara, K., and Scerri, P. (2006). Scaling-up human control for large UAV teams. *Hum. Factors Remot. Operat. Vehicles* 7, 237–250. doi: 10.1016/S1479-3601(05)07017-7

Mekdeci, B., and Cummings, M. L. (2009). "Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles," in *Proceedings of the 9th Workshop on Performance Metrics for Intelligent Systems - PerMIS 09, Association for Computing Machinery* (Gaithersburg, MD). doi: 10.1145/1865909.1865911

Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. doi: 10.1016/S1364-6613(03)00028-7

Mosier, K. L., and Skitka, L. J. (1996). "Human decision makers and automated decision aids: made for each other?," in *Automation and Human Performance: Theory and Applications*, eds R. Parasuraman and M. Mouloua (Mahwah, NJ: Lawrence Erlbaum Associates), 201–220.

NHS England (October 2014). *Five Year Forward View*. Available online at: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (2004). *Defense Science Board Study on Unmanned Aerial Vehicles and Uninhabited Combat Aerial Vehicles*. Washington, DC: Department of Defense.

Oron-Gilad, T., Porat, T., Fern, L., Draper, M., Shively, J., Silbiger, J., et al. (2011). "Tools and techniques for MOMU (Multiple Operator Multiple UAV) environments; an operational perspective," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 55, 86–90. doi: 10.1177/1071181311551018

Parasuraman, R., Kidwell, B., Olmstead, R., and Lin, M.-K. (2014). Interactive effects of the COMT gene and training on individual differences in supervisory control of unmanned vehicles. *Hum. Factors* 56, 760–771. doi: 10.1177/0018720813510736

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybern.* 30, 286–297. doi: 10.1109/3468.844354

Porat, T., Oron-Gilad, T., Silbiger, J., and Rottem-Hovev, M. (2010). "'Castling Rays' a decision support tool for UAV-switching tasks," in *CHI 2010 Conference Proceedings* (Atlanta). doi: 10.1145/1753846.1754023

Porat, T., Oron-Gilad, T., Silbiger, J., and Rottem-Hovev, M. (2011). "Switch and deliver: display layouts for MOMV (Multiple Operators Multiple Video feeds) environments," in *IEEE CogSIMA 2011 Conference Proceedings* (Miami, FL). doi: 10.1109/cogsima.2011.5753457

Ruff, H. A., Calhoun, G. L., Draper, M. H., Fontejon, J. V., and Guilfoos, B. J. (2004). "Exploring automation issues in supervisory control of multiple UAVs," in *Presented at 2nd Human Performance, Situation Awareness, and Automation Conference (HPSAA II)* (Dayton Beach, FL).

Ruff, H. A., Narayanan, S., and Draper, M. H. (2002). Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles. *Presence Teleoper. Virtual Environ.* 11, 335–351. doi: 10.1162/105474602760204264

Saqer, H., de Visser, E., Emfield, A., Shaw, T., and Parasuraman, R. (2011). "Adaptive automation to improve human performance in supervision of multiple uninhabited aerial vehicles: individual markers of performance," in *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (Las Vegas). doi: 10.1177/1071181311551185

Shaferman, V., and Shima, T. (2009). "Task assignment and motion planning for multiple UAVs tracking multiple targets in urban environments," in *AIAA Guidance, Navigation, and Control Conference* (Chicago, IL). doi: 10.2514/6.2009-5778

Sheridan, T. B., and Verplank, W. (1978). *Human and Computer Control of Undersea Teleoperators*. Cambridge, MA: Man-Machine Systems Laboratory; Department of Mechanical Engineering, MIT.

Squire, P. N., and Parasuraman, R. (2010). Effects of automation and task load on task switching during human supervision of multiple semi-autonomous robots in a dynamic environment. *Ergonomics* 53, 951–961. doi: 10.1080/00140139.2010. 489969

Vaishnavi, V., and Kuechler, W. (2004). *Design Science Research in Information Systems*. Available online at: http://www.desrist.org/desrist

Voshell, M., Woods, D. D., and Phillips, F. (2005). "Overcoming the keyhole in human-robot coordination: simulation and evaluation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 49, 442–446. doi: 10.1177/1541931205049 00348

Wilson, G. F., and Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Hum. Factors* 49, 1005–1018. doi: 10.1518/001872007X249875

# "If It Feels Right, Do It": Intuitive Decision Making in a Sample of High-Level Sport Coaches

Dave Collins[1,2]*, Loel Collins[1] and Howie J. Carson[1]

[1] Institute for Coaching and Performance, University of Central Lancashire, Preston, UK, [2] Grey Matters Performance Ltd., Stratford upon Avon, UK

Comprehensive understanding and application of decision making is important for the professional practice and status of sports coaches. Accordingly, building on a strong work base exploring the use of professional judgment and decision making (PJDM) in sport, we report a preliminary investigation into uses of intuition by high-level coaches. Two contrasting groups of high-level coaches from adventure sports ($n = 10$) and rugby union ($n = 8$), were interviewed on their experiences of using intuitive and deliberative decision making styles, the source of these skills, and the interaction between the two. Participants reported similarly high levels of usage to other professions. Interaction between the two styles was apparent to varying degrees, while the role of experience was seen as an important precursor to greater intuitive practice and employment. Initially intuitive then deliberate decision making was a particular feature, offering participants an immediate check on the accuracy and validity of the decision. Integration of these data with the extant literature and implications for practice are discussed.

Keywords: adventure sports, coaching practice, expertise, macro cognition, professional judgment and decision making, rugby

## INTRODUCTION

Intuition is of increasing interest to expertise researchers (e.g., Dane et al., 2012). Certainly, in the present context there is considerable anecdotal evidence that sports coaches often prefer to "go with the gut" (cf. Lyle and Cushion, 2010), taking fast action on the basis of what feels right (known as Naturalistic Decision Making; NDM) rather than through a more formal and slower Classical Decision Making (CDM; see Abraham and Collins, 2011) style reasoning which explicitly balances the options. Independently of how intuition is seen to operate (and various theories do offer different perspectives on the mechanism; cf. Klein, 2015), we were interested to build from our recent work in coach decision making (DM; e.g., Collins and Collins, 2012, 2013, 2015a,b, 2016b) to examine the role that intuitive decisions were perceived as playing in the repertoire of high-level coaches. Reflecting our interests and experience, and also to offer a contrasting pair of environments, we decided to conduct this primary exploration in adventure sports and rugby. Use of these settings enabled comparison of DM between a hyper-dynamic, perceived high-risk environment (adventure sports or AS; Collins and Collins, 2013) and a more conventional, less time pressured, and more ego than physical risk situation (at least for the coach!) such as the team sport of rugby union (RU).

Defined as "the capability to act or decide appropriately without deliberately and consciously balancing alternatives, and without following a certain rule or routine, and, possibly, without

awareness" (Harteis and Billett, 2013, p. 146), intuition offers a significant extension to the concept of "on-action, in context" reflection described by Collins and Collins (2015a, p. 622) in their work with AS Coaches. Although, certainly with awareness and involving some balancing of alternatives, the on action-in context approach suggests an immediate review or "audit" of a quickly taken decision, offering the coach both a safety check and a reassurance that the gut feel action is correct.

The two environments used in this paper offer parallel but contrasting challenges for coach DM. The coaching of AS has emerged as a sub-set of coaching practice that draws on knowledge from the sports coaching domain and outdoor education (Collins and Collins, 2012). AS coaching practice is characterized by a high cognitive load brought about by the challenges associated with a hyper-dynamic coaching environment (Collins and Collins, 2016a). Specifically, the need to respond to challenges of a literally relentless, constantly changing physical environment, the infinite nature of the learners, and the relationship of the two provides a unique "wicked" challenge (Horn and Weber, 2007) to the coaching process. In our paper, examination of DM was considered across environments, but mostly focused on the majority environment, the teaching and development of clients in the AS situation, be it on the mountain, the river or the sea (cf. Collins and Collins, 2012).

In the more usual coaching environment of a team sport like RU, intuition is also an attractive construct. Decisions taken by high level coaches in the professional sport environment are certainly high stakes. Indeed, the high levels of challenge and turnover apparent (Cruickshank et al., 2014) make such environments equally high risk, albeit in a different way. In the data presented in the present paper, although once again all elements of DM were examined, comparisons were often drawn between teaching/training environments which, for RU coaches, also represent the majority of their coaching work with players. Environmental differences notwithstanding however (and these were carefully considered in our analysis), both these coaching settings impose high levels of challenge on coach–practitioners, offering an effective test-bed for our preliminary investigation of intuition in coach DM.

For practitioners in both, the implied expertise associated with intuitive or tacit DM has already been mentioned and even promoted in the coaching literature (cf. Lyle and Cushion, 2010). As such, there would seem to be some significant advantages to the adoption of intuitive DM for coaches, paralleling those already shown for performers (e.g., Janelle and Hillman, 2003; Raab and Laborde, 2011), so long as these were shown to generate decisions of equal (or even better) accuracy. Clearly, quicker and less effortful processing represents one big advantage: the possession of a knowledge base of sufficient richness to support/encourage intuitive DM is another concomitant benefit. Quite apart from the balance of advantage and disadvantage, however, there is a real need to achieve a more comprehensive understanding of coach DM.

In this regard, it is important to acknowledge the importance of research into professional judgment and decision making (PJDM) for the development of coaching practice and its status

as a profession. Clearly, both the environments explored in this paper offer high stakes DM, with pressures varying from life and death (AS) through to professional rewards for success and sanction for failure (RU). Furthermore, understanding and effectively utilizing the cognitive and macro cognitive skills of coaches as the basis for assessment and professional development is increasingly recognized (cf. Collins and Collins, 2016b). In short, a move from a simple behavioral competency model to one firmly based in expertise is both overdue and would offer a significant step forward for coaching (see Collins et al., 2015). Certainly, the use of this approach would represent a definite move toward the recommendations associated with competence in professional settings (Kaslow et al., 2007). As such, examination of coach DM is of extreme utility to the profession, as well as an important topic in its own right.

In fact, research to date suggests that DM may take place on a continuum between CDM and NDM, with intuitive DM lying even further along the NDM end. The concept of nested DM (cf. Martindale and Collins, 2007, 2012; Abraham and Collins, 2011; Collins and Collins, 2015a,b, 2016b) as a part of the application of PJDM (op cit) to coaching, saw higher-order/longer-term decisions as best taken in a more considered deliberative (CDM) fashion while immediate, in-session decisions were more short-term and almost intuitive (more reflective of an NDM approach). The nesting of the latter within the former, so that short-term decisions generally took into account and catered for longer-term agendas, was suggestive that intuitive DM in a predominantly cognitive task such as coaching may show such an interaction.

Accordingly, and reflecting these different perspectives, we were interested to examine several pertinent aspects of coach DM as follows:

- the nature, scope, and incidence of intuitive DM in samples of high-level coaches;
- the relative frequency and origin of this DM style, at least as perceived by the sample, and;
- to what extent intuitive decisions interacted with more deliberative DM.

## METHODS

### Participants

Participants were 18 male British coaches from rugby union (RU: $n = 8$; $M_{age} = 48.2 \pm 3.3$ years) and adventure sport (AS: $n = 10$; $M_{age} = 43.5 \pm 12.5$ years) domains. To ensure a sufficient level of domain expertise, experience, and inherent quality in terms of participants' self-reflective ability, purposive sampling was employed based on the following criteria: (1) a minimum of 10 years coaching experience since senior accreditation (RU: $M = 14.9$ years; AS: $M = 15.1$ years), (2) currently working with internationally-competitive and/or higher (e.g., professional/premiership) performers and/or hold the highest level coaching qualification within their respective sport, and (3) have a willingness to discuss their professional practice. All of the coaches were recruited through personal contact with the research team; the corresponding and second authors here being qualified and active practitioners within

these two respective high-level sporting domains. This study was carried out in accordance with the recommendations of University of Central Lancashire's ethics committee with written informed consent from all participants in accordance with the Declaration of Helsinki.

## Procedure

To enable sufficient breath and richness of responses to be explored, a qualitative methodology was adopted. Specifically, semi-structured interviews were conducted with each coach in a quiet, private location, and at a time convenient to them. Participants received an information sheet by email at least 1 week prior to interview and, after consenting, the interview commenced by flexibly covering the lines of questioning shown in **Table 1**. In brief, the interview guide asked participants to recall and evaluate coaching episodes where DM utilized careful thinking (i.e., deliberative, CDM style) and others through sudden insight (i.e., intuitive, NDM style). In designing the questions, we were informed and guided by the work of Crandall and Getchell-Reiter (1993) whose application of the Critical Decision Method to nursing incidents in critical care offered a strong template. The classic and naturalistic types of DM were also explored more generally, as too were the learning experiences of each participant, and perceived skills, and attributes required to improve one's DM efficacy. Probes were deployed where necessary to gain additional information relating to interesting/important responses and/or check ideas against emerging literature, thus ensuring sufficient depth of response across all participants.

Two researchers conducted the interviews and analysis of corresponding transcripts (see below), both are highly experienced in their respective fields and therefore were able to question, probe, and interpret responses with a degree of seniority. One of the researchers has 30 years of experience as an ASC at National Centers within the United Kingdom, is a coach educator, and holds Level 5 British Canoe Union coaching awards in four disciplines. The other researcher holds senior coaching qualifications in rugby, has experience of national level coaching in the United Kingdom and abroad, and has worked as a support professional in rugby at international level. Overall, the entire interview process lasted between 60 and 90 min. Data were recorded using a Dictaphone and stored electronically in mp3 file format.

## Data Processing and Analysis

Following the guidance provided by Braun and Clarke (2006), data were analyzed using a thematic analysis. Accordingly, interviews were first transcribed verbatim and each transcription was actively read several times prior to fully apprehend the essential features (Sandelowski, 1995). General impressions of these data were written in note form and shared between the two researchers conducting the analysis (first and second authors), highlighting any similarities and differences. Secondly, driven by an analytic interest of DM processes and informed by the literature, initial inductive coding of response data was applied to each transcript; thus formally identifying relevant and similar extracts. Thirdly, data codes were collated into potential

lower-order themes based on common features, which were then grouped together under higher-order themes representing the highest level of abstraction. Within a fourth phase of analysis, these themes were subjected to review and further refinement. Meetings were held between the two researchers to discuss and compare the analysis between rugby and adventure sport domains. The primary aim was to check for a shared understanding and interpretation of data and, therefore, the emerging themes as a whole data set. This process was essential to detect genuine effective equivalence between situationally-specific behaviors; clearly the two groups were looking to generate rather different outcomes. As such, it was both informative and interesting that a high degree of overlap occurred. Our approach enabled themes to be combined and broken down, as well as the generation of new themes. Importantly, and reflecting our desire to examine genuine rather than artifactual (or even investigator created) overlap, the development of themes at any point during the analysis did not depend on the prevalence of a code, but rather, on what the theme revealed about the DM process. Finally, again as a collaborative process, the two researchers defined themes according to the essence of data codes within and how these might be perceived in relation to other existing theme definitions against the particular context of adventure sports or rugby.

In addition to the steps outlined above to ensure inter-coder agreement, the issue of trustworthiness was addressed through use of an additional researcher, who was not involved in the interviewing or coding process, independently coding a random sample of the transcripts (20%). This researcher coded raw data against the developed themes and his results were compared to those derived by the original process. Any disagreements (four emerged) regarding these differences in codes were discussed until a consensus was reached. This researcher also examined the overlap across domains, through interrogation of the equivalence derived from the thematic tables. Once again, disagreements (three emerged) were debated until a consensus was reached. Importantly, almost all of these disagreements fell within the first higher-order theme of learning environment.

## RESULTS

A breakdown of the parallel thematic analyses are presented in Appendices A and B in Supplementary Material, with a summary table (**Table 2**) to exemplify points of equivalence and situationally specific difference. Unsurprisingly, differences were most apparent in the more situationally-specific settings of learning environment, although even here there was a very high degree of equivalence, most apparent in the intermediate, and higher-order themes arrived at by consensus across the research team. Reflecting the stated objectives of this paper, we report on a subset of the data yielded by the investigation, which predominantly draw on the second and third higher-order themes, which focus most specifically on our purpose. Following the structure outlined earlier, we report participant views relating to the three main objectives, followed by other relevant material from the interviews relating to the

**TABLE 1 | Interview Guide.**

| Question | Probes | Stimuli | Purpose |
|---|---|---|---|
| 1. Can you tell me about a recent coaching episode when you came up with a novel solution to a problem through careful thought? | • What did you do and with what aims?<br><br>• What did you believe were the key parts for achieving these aims? | • Performance improvement<br>• Athlete/team development<br>• ID of the problem<br>• Generating and considering alternatives<br>• Experimentation?<br>• Criteria for PJDM | • Explores incidence of reflective practice and "insightful experimentation" |
| 2. Can you tell me about a recent coaching episode when you came up with a novel solution to a problem through sudden insight? | • What did you do and with what aims?<br><br>• Can you post hoc rationalize where the solution came from?<br><br>• Can you think of any learning, knowledge or training which helped you generate this solution? | • Performance improvement<br>• Athlete/team development<br>• Just popped into your head?<br>• Whilst coaching or away from the context?<br>• Something which was on your mind for a while?<br>• Recent or past training, reading or viewing?<br>• Discussions with peers | • Explores incidence of tacit knowledge |
| 3. To what extent are your coaching decisions intuitive or carefully planned?<br>HOW and WHY? | • Percentage incidence and context for each<br>• Personal preference<br>• Which do they think are most effective and why? | • Try for a numeric balance across 100%<br>• Which do they prefer and why?<br>• Measure of efficacy they use? | • Personal balance of styles |
| 4. Where do you usually gain knowledge on coaching?<br><br>Can you relate this to a recent change which you have made? | • Formal/informal balance<br><br>• What is good about preferred approach<br><br>• What is bad about non-preferred approach?<br>• Have your preferences changed over time? | • Highlight previous successful learning<br>• Fits personal style<br>• The absolute best against theory<br>• In my experience<br>• Inverse of above<br>• Younger versus now preference and why? | • Explores sources of learning |
| 5. Can you tell me about any personal attributes and skills that you believe were important for carrying out the changes? | • Intelligence<br>• Open mindedness<br>• Critical reflection<br>• Adaptability<br>• Recognizing pivotal movements in the process | | • What does it take to be an innovator? |

evolutions underpinning the use of intuition. For clarity and confidentiality, coaches are referred to by sport (RU or AS) and a number.

## The Nature, Scope, and Incidence of Intuitive DM

All coaches recognized situations in which intuitive decisions were apparent in their practice. Coach AS2 articulated this as follows:

> I think much more on my feet now. I'm much more intuitive as a coach than I ever was before. I come out with less and less

structured [pre-planned] sessions. I've got structured sessions in the back of my mind, they're ingrained there, I've done them over and over again but what I do now is…I'll, kind of…I will adapt that. I will adapt that [the session] not only to the situation but also to the mood [the group's responses to his coaching] and to the environment as well.

This was echoed by coach AS9 who highlighted:

> I prefer the one [DM style] that is right for the situation I'm in…I'd probably start with logical and linear, but there is a hell of a lot more intuitive that appears and it tends to go that way, because the likely course with a client, especially if it's

TABLE 2 | Summary comparison of key themes for adventure sport and rugby coaches.

| Higher-order theme | Mid-order theme | Adventure sport coaches | Rugby union coaches |
|---|---|---|---|
| Learning environment | Pedagogic context | 10/10 | 8/8 |
| | Environmental context | 10/10 | – |
| | Season context | – | 8/8 |
| Experience | Professional | 10/10 | 8/8 |
| | Reflective skills | 10/10 | 8/8 |
| Meta process | Aspects of decision making process | 10/10 | 8/8 |
| | Audit (change of perspective on DM) | 8/10 | 8/8 |

long-term [characterizing the student's relationship with the coach], you're less likely to be following a logical linear path after a while. You're more likely to be reacting to what's going on around you, their development rate…So the linear journey to a goal might fall apart, especially in the long-term.

Rugby coaches were similarly universal in reporting that intuitive decisions fitted their behavior in certain settings. RU3 said "well I will take a gut feel and apply it, but almost always during a session or game." RU6 added "we are surrounded with so much data on game day; but there is still a real part for intuition on substitutions, tactics, and the like." Furthermore, all coaches acknowledged a role for intuition in player signing and selection. As Coach RU7 forcibly stated:

We get so much detail about what this guy is like…from agents, scouts, committee; f***, everybody gets in on the act. But at the end of the day, it's my call. And I make it almost entirely on feel…would I like to play with him, does he fit the [Club name] tradition, is he a good bloke as well as a player.

The intuitive characteristics of the DM process reported by participants appeared tacit (difficult to articulate) in nature but also to be based on refined and integrated reflective practice of a long, varied experience. As such, they emerged as aspects of a micro process, to meet a short-term challenge, but emanated from a longer-term macro process of development. Several coaches encapsulated meeting the challenges through intuitive decisions, but also highlighted the role of previous reflective practice as conferring an ability to *ad-hoc* rationalize a significant decision and unpack tacit aspects of the knowledge once utilized. AS5 described the decisions associated with observation of a group of whitewater kayakers prior to a session:

A lot of them [decisions] are sub…almost subconscious, that I don't quite know…, Like…, I don't sit there and go, right I've got five options that I could do with these people. It's more watching them on the water and thinking, well, what are they [doing].

AS4 encapsulated the gut feel as "I think it's how comfortable I am about sorting it out" when discussing a safety specific point. AS7 used similar descriptions of his decisions in relation to teaching and allowing the students to "…just let it happen, because I wanted to see how they [the students] would perform." In both

cases, the "gut feel" was to gain benefit (learning) from a risky situation (cf. an intuitive risk–benefit decision; cf. Collins and Collins, 2013). Both were confident that they could "sort out" the consequences so let things progress because "it felt right."

Rugby coaches were identical in their reports of intuitive feel around running training sessions. RU2 spoke of how he would "let things run over, or change direction totally, going away from the plan to take a new direction just because it feels right." Several spoke of how shared models and intuition with established colleagues enabled them to take new and novel directions in sessions with minimal or no discussion. RU4 reported on his long-term working relationship with another coach: "we just look across at each other, often without warning, nod and just both start working on something new; often in an entirely unplanned direction." RU6 explained "I've been doing this for a long time…I think I coach now like a musician in a jam session!"

All these quotes are illustrative of a trait common to *all* the coaches interviewed; knowledge made usable and reliable in context by it becoming tacit following a period of reflection on extensive experience. The types of rule which emerged from this *post-hoc* reflection generates powerful tools for the future. Certainly, the use of rule-based strategies has been shown as a good way to handle unexpected or novel challenges (Richters et al., 2015). In this case it is interesting to see the ways in which rules may have emerged from initially intuitive action. Notably, however, this internalization or automation of earlier decisions, whether taken intuitively or through a more classical process, seems to occur through reflective processes; thinking through and weighing up the action before it is accepted as useful and locked into that individual's repertoire for subsequent, more intuitively led employment.

## The Relative Frequency and Origins of Intuitive DM

There were considerable variations, both between participants and between sports on the perceived frequency of intuitive DM. In all cases, participants acknowledged the need for careful planning across all elements of their work. Interestingly, however, the intuitive aspects of the coaches' DM emerged differentially across the macro and micro processes of the session. Within the macro process this could be observed in the planning stages of a

session, and within the on-action/in context decisions that took place in the time generated for thinking within the session by the coach (cf. Collins and Collins, 2015a). Coach AS6 encapsulated the dilemma for the AS coach: "I suppose it would become intuitive if it's becoming something that you can't control from planning."

For several, intuition seemed to be a feature of personal preference and professional context. Coach AS10 stated, "I think far more of it's rational than intuitive, but that's because I'm…. because of the nature of what I'm doing." This contrasted with Coach AS5 who appears to state the opposite by explaining, "I'd say almost 90, probably 90% intuitive, 10% working out." All AS coaches found difficulty in allocating a percentage to their DM process and qualified their original percentages when pushed. Coach AS5 qualified his original estimate by following his original quote (above) directly with "I'd say the working out was just at the beginning of the session."

In contrast, rugby coaches were both more consistent and, perhaps, more conservative in their estimates of intuitive DM. As a typical comment, coach RU8 (the most senior and experienced of the sample) stated:

> we tend to be quite rational and careful in planning…in thinking things through and justifying actions. Perhaps that is the team thing; we have to sell it to the coaching team and justify it to the players, especially when they are senior professionals. However, even the most staid coaches will, in my experience, take a leap in the dark sometimes; on a player, a substitution, a move or a change of plan.

As a consequence, perhaps, RU estimates of intuition DM percentage were on average lower, around 30%, with a much smaller range of 5–40%.

With regard to origin, participants were unanimous in acknowledging that their effective use of intuitive style had come with experience (cf. Pretz and Folse, 2011). Highlighted by Harteis and Billett (2013) as "the common elements of highly learnt procedures and informed strategic capacities that, together, support the capacity to act intuitively and with great effect" (p. 146), there seemed little doubt for our sample that they could now perform intuitively only because of a long and rigorous apprenticeship. Coach AS9 stated, "it's applying that decision making process in lots and lots of different situations over lots and lots of years in my case." In similar fashion, RU4 explained:

> when I started, I planned meticulously and almost agonized over decisions in case I got them wrong: but now I just go for it…I'm secure in my experience and can judge the situation as something I've seen before—so I can fly faster and with less thought—or something new which needs more careful thought.

This relationship, between situational awareness and the decision makers' experience and skill, contains the interactional aspect of the process already highlighted in our earlier work (Collins and Collins, 2016b).

The roles played by others, such as coaching companions, team members or, as stressed by the AS coaches, their community of practice (CoP), seemed particularly important in helping this group to the knowledge levels and associated confidence necessary for effective deployment of an intuitive style. Extending his ideas above, RU4 explained "I would have to credit my fellow coaches, my mentors, in building the knowledge and confidence which helped me "loosen up" and get intuitive." All the rugby participants highlighted senior coaches (described as, but never formally in, a mentor role) as a major source of coaching knowledge and the support to make changes based on intuition; "to go with the gut" (RU3). Notably, however, self-directed reflective practice as a consistent strategy was far less common, with only two RU coaches describing this as an explicit and regular part of their development repertoire. CoPs, where reported, were perhaps understandably restricted, usually within the participants' club or to particular friends in the field.

In contrast, AS coaches all reported a high degree of personal reflective practice and engagement with their CoP: the role of the CoP being as a critical friend, a sounding board, and an exchange of coaching related knowledge. Coach AS10 highlighted "I think I'm lucky to be part of a community of paddlers" and further explained "I get to see and hear other peoples' [coaches] perspectives." Coach AS4 described the characteristic of a productive CoP working in mountainous artic conditions; "it's very much a supportive culture, people are quite happy to ask about advice and some is a bit wacky [the advice] and some will say, I don't think you should go there today." Coach AS5 described the characteristics of an effective CoP with a coaching focus; "Yes. Yes, definitely. Yes, just willingness for everyone to go, "oh, you do it that way and you do it that way," or not…not having any fixed…fixed way of doing it." Notably, the seven AS participants involved in coach and leadership education, in addition to their skills development role, were able to articulate and clearly value their DM knowledge and skills originating from regular CoP interactions.

## The Interaction of Intuitive and Other DM Styles

As intimated in several of the quotes above, participants were very aware of the parallel and/or interactive use of intuition with other, more deliberative styles of thinking in their coaching. Interestingly, they reported very few examples of intuition as a quick "that will do" but suboptimum alternative (cf. fast and frugal; Gigerenzer and Todd, 1998). All the coaches in this study recognized both intuitive and classic characteristics in their DM processes during coaching. A sizable minority treated the two as somewhat distinct, suitable for use in certain circumstances. As RU2 reported:

> I think it reflects my original profession [uniformed services]. I recognize a situation as requiring decisive decisions and get myself in the headset to act so. This almost always includes making big calls on feel. I do debrief them later but, in the moment, its card laid, card played!

In similar fashion, coach AS6 stated "I think if done right they're both effective," referring to considered and intuitive processes although implying a non-nested relationship. Coach AS7 linked the characteristic of the DM process to other aspects of practice and clearly illustrated a comprehension of the characteristics of his DM process:

> I would like to say, when I was learning and gathering my experience, it's definitely that planned [considered] approach because you felt safe, you felt okay…Whereas now [referring to intuitive characteristics], I kind of, almost get a bit more excited, I didn't expect that, that's great. Let's go with it and see how it goes, probably because you know at the back of your head, that if it starts to go wrong, you can still fix it and put it back on track.

This appears to illustrate a necessary confidence in the intuitive elements that has emerged over time and practice. These coaches declared personal preferences for a given approach, with the preference reflecting their experience and personality; importantly, the coaches all articulated a confidence in being more intuitive or considered because of an imbedded audit of the decision making process.

Going further, however, a majority (8/10 AS and 4/8 RU) demonstrated an additional meta-process in the more integrated use of the different styles, based around a clear recognition of the particular advantages or disadvantages in that context. This idea resonates with the "rich systematic interactions" identified by Christensen et al. (2016, p. 40) as crucial between automaticity and cognition in movement execution. Under this "Mesh" approach, athletes exhibit a delicate but consistent balance between cognitive and automatic elements of control, except when the balance is disrupted (often by anxiety) toward an overly cognitive style. The relationship between the intuitive and classic aspects of DM in our participants was nested in nature and influenced by two factors; a context-based, situational awareness and the decision maker's experience and skill. For example, AS5 suggested he chooses from "millions of options" and recognizes "I do like a coaching problem." In all these cases, the DM process, including a refined reflective practice, was imbedded within the coaching process and audited by the coach.

These 12 coaches articulated an *ad-hoc* triangulation/audit of the DM process that was achieved via a notional question: "is the same outcome achieved via a different DM approach?" The audit was used to verify or challenge the original decision which informed action on a particular course of action. This triangulation/audit is time consuming and adds to the time pressures but was considered valuable given the complexity of the environment and consequences in sport. This audit was integrated into the process by creating time for DM, not just for the original decision but also for the audit (Collins and Collins, 2015a). As RU1 observed "I decide to do something, say make a substitution, but immediately I'm scanning the decision to see if it feels right." These descriptions fit well with the parallel systems ideas of Myers (2002) and Sadler-Smith (2010), representing the twin use of intuition and deliberation to generate optimum solutions.

Interestingly, however, and perhaps representing an extension to the parallel systems ideas, this audit process did not necessarily use the alternative deliberative style but was often also intuitive in its nature. Often, the use of intuition to audit intuition was determined by pressure; from the environment, context or, most notably, emotions (cf. Slovic et al., 2002). It was in these situations that intuitive skills appeared most valued by the coaches. AS4 described a forced decent from a winter climbing route in deteriorating conditions in which his decision "was the least of all bad options, there weren't many…there was no good option and it was the least of all the bad options, there were no good options really." He later described the decisions as needing to "go with your gut" (the primary, let's retreat decision) while asking a rhetorical question of himself "does this feel right" as the auditing process for the route selected. In similar fashion, RU5 reported "in that situation I felt really angry. I wanted to take action so made the call, at the same time thinking to myself "does this feel right"?"

The work of Eraut (1994, 2000) offers a very parsimonious explanation of our data on style integration. As he states, intuitive responses may be represented as:

> not only pattern recognition but also rapid responses to developing situations... based on the tacit application of tacit rules. These rules may not be explicit or capable of reasoned justification, but their distinctive feature is that of being tacit at the moment of use (Eraut, 2000, p. 127).

AS4 articulated this challenge in needing to rely on easily accessible decision making skills and demonstrates a need for confidence in the NDM process together with a realization of a meta-process that exists within the NDM aspects of the process. AS4 was torn between gut feel and recognition primed DM, and a capacity to articulate the complexity to his students:

> ….trying to convince the students [articulating the dilemma] that that was a really serious day and the decision making, they all thought it was fantastic and it was a really exciting adventure, but you know, it's trying then to tell the students, actually…there was some wrong decision making going on there. There was some gut decision making that was…that basically was fine up to a certain stage [limitations of a given approach], but then it's the conditions and the environment changed [situational awareness, change, and impact] and so I was stuck, having to make gut decisions [other processes may have been better suited], and realizing that I was now in a situation that wasn't good [audit].

*Post-hoc* recall and rationalization of decisions, however unconscious/intuitive, was a common feature across participants. As RU1 observed:

> In the heat of battle, I say and do all sorts of things. My coaching team and analysts often look at me strangely to think "why the f*** has he done that?" But I can always run the replay in my head afterwards, with total recall, and explain the logic of why, when, and how even though I wasn't aware of it at the time.

## DISCUSSION

While the existence and use of intuition as a DM approach was clear from our data, when and how it was used is of greater interest, especially when contrasted to the application of more deliberative strategies. Clearly, the proportion of considered to intuitive characteristics in a given decision varied depending on context, based largely on the coach's high situational awareness of a given session or context (Endsley, 1995a,b). This awareness seemed to be situated within practically set parameters that framed the process, and included elements that are managed by the coach (e.g., logistics, equipment, student/player preferences, fitness, cognitive ability) plus, potentially, learning outcomes, (i.e., an outline plan), and some that cannot be directly managed (e.g., tides, snow conditions, weather or player reactions, game outcomes, etc.). The role of the *post-hoc* audit check (whether quick or more deliberate) is another crucial finding for future work; intuitive decisions in coaching may not be as unaware as Harteis and Billett (2013) suggest. Such an action is understandable in the hyper-dynamic, high stakes environment of AS. Finding it in the more conventional and less (comparatively) time pressured world of mainstream RU coaching (the majority of time, and our data, came from the training environment) however is, perhaps, more surprising. This idea merits more detailed examination.

Extending this interactive theme, it is worth considering more carefully how the relationship between deliberate and intuitive DM may operate. This relationship is clearly "perceptual" in its nature and operates continually within the coaching process, forming the coaches' awareness to the situation (situational awareness) as it evolves. In the present sample, this appeared to be related to three interacting contexts. The first, the pedagogic context, appears to be comprised of a further set of sub-factors such as the learning outcomes, syllabus/issue content, potential goals, and the nature of the individuals being taught. The second environmental context relates to the real and perceived risk to the participants by considering the physical and social environments of the decisions. In rugby, though less serious, this session context is also clearly important. These two, the pedagogic and environmental/situational content, interact to form a third subgroup, the learning environment, which links to the decision maker's experience and skill in DM. Our suggestion here is that effective use of an intuitive DM style, indeed all styles, will be determined by the education which the coach receives on how coaching works (cf. theories of knowledge generation; Nonaka and Takeuchi, 1995). Accordingly, and particularly from an evolutionary perspective, coach educators may also need to consider the environment in which such evolutions may be optimized (e.g., Nonaka and Nishiguchi, 2001).

Irrespective of how they may best be developed, intuition and analysis are both important components of expertise and their mutuality seems well supported by our data (cf. Pretz, 2011). Interestingly, this interplay seems to be most important in certain settings and conditions. In AS coaches, for example, the interplay of considered and intuitive characteristics form part of the coaches' ability to rapidly adapt and be flexible in session, this appears to be facilitated by the intuitive aspects of the process. The in-action decision process that requires greater flexibility and adaptability having more intuitive characteristics. Adaptability and flexibility (modification of existing knowledge) was considered by all the AS coaches as an aspect of their expertise. A lower number of AS coaches ($n = 7$) identified creativity (creation of novel solutions in-action or on-action/in context) as an aspect of their expertise. In this respect, our own rule for the use of new information in a coaching situation involves an immediate scan and placement as follows…"Act on, Store or Ignore": in short, making a rapid initial evaluation on the potential worth of new data. This enables the essential rapid action within the hyper-dynamic environment of adventure sports; without it, the coach may quite literally drown in data and be paralyzed by in-action reflection.

The position of rugby coaches was somewhat different on the first aspect, most particularly because their DM was usually less time pressured and dynamic than AS. As such, the interaction with deliberative DM was more marked and frequent. There was much greater concordance on creativity, however. Almost all ($n = 7$) of the RU coaches mentioned the role of intuition in generating novel and creative solutions to problems they encountered. Clearly, the aspect is particularly worthy of examination, especially at the top end of performance where originality is often key to success.

Another important issue for the future relates to how intuition might best be investigated. We acknowledge the comparative crudeness of our "percentage of intuition" question (Q3 **Table 1**); also the limitations inherent due to small numbers in our sample, albeit these participants are of a high level and hence, drawn from a small population. Our point is the self-reported difficulty participants experienced in answering the question, which means that future studies must use tracking to generate more accurate figures. We see the present study as a first exploration and, as such, report the data accordingly. For the future, however, in our own and other's work, how this tracking is best accomplished is an issue. It is not, we suggest, just a case of "think out loud" (cf. Whitehead et al., 2015) although our data suggest that an immediate internal or even external audit often follows a gut feel decision. Perhaps the best option is a more naturalistic observation of the process, with immediate follow up and critical probing to take the participant back through the situation soon after it has occurred. The instance-driven interview questions used in this study (see **Table 1**) are an NDM research technique that can/should be used to operationalize Intuitive DM in ways that can be easily interpreted and subsequently included in coach education initiatives, so long as the *post-hoc* elements described are also tested for. We have certainly used this approach to good effect in our own work on coach DM (e.g., Collins and Collins,

2015a,b, 2016b). In any case, methodological issues will be an important consideration for the future.

We also need to investigate the evolution of intuition in coaching, especially since its popularity and high credibility status (cf. Tetlock, 2006) may "encourage" such habits in beginner coaches. After all, peoples' preference for decisiveness has been well documented! Our participants were certainly supportive of the conventional wisdom that intuitive DM emerges from experience. Once again, however, research from nursing offers some interesting parallels and contrasts. For example, Ruth-Sahd and Tisdell (2007) suggest that the use of intuition is more related to previous experience with it as a style than level of training. Others see the use of intuition as a trait (e.g., Myers et al., 1998; Pacini and Epstein, 1999). These issues notwithstanding, there seems to be a strong case for intuition as a characteristic which emerges from experience (cf. Pretz and Folse, 2011) and, as such, more advanced training may well-include its use as a consideration using suggestions from the NDM literature as a basis (cf. Klein, 2015).

There is one other issue worthy of note; that is, against the definition used of intuition as an unconscious act, our data seem to suggest that the process (at least as seen by our coaches) is often semi-conscious or, even if unconscious at the time, almost immediately brought into the conscious space and rapidly reviewed. Certainly, several of the studies cited in this paper have highlighted this conundrum (most notably the nursing research). Is there, perhaps, the need for a new model with regard to the definition of intuitive thinking? We see interesting and important parallels between these ideas and the Mesh control suggested by Christensen et al. (2016) as a parsimonious solution to the interplay of conscious and automatic processes in movement. This issue awaits further examination. For the moment, however, the place of intuition in the DM of high-level coaches is clearly established, albeit that it might be less automatic and implicit than some popularist authors may suggest.

## AUTHOR CONTRIBUTIONS

## SUPPLEMENTARY MATERIAL

## REFERENCES

Abraham, A., and Collins, D. (2011). Taking the next step: ways forward for coaching science. *Quest* 63, 366–384. doi: 10.1080/00336297.2011.10483687

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Christensen, W., Sutton, J., and McIlwain, D. J. F. (2016). Cognition in skilled action: meshed control and the varieties of skill experience. *Mind Lang. (Baltim).* 31, 37–66. doi: 10.1111/mila.12094

Collins, D., Burke, V., Martindale, A., and Cruickshank, A. (2015). The illusion of competency versus the desirability of expertise: seeking a common standard for support professions in sport. *Sports Med.* 45, 1–7. doi: 10.1007/s40279-014-0251-1

Collins, L., and Collins, D. (2012). Conceptualizing the adventure-sports coach. *J. Advent. Educ. Outdoor Learn.* 12, 81–93. doi: 10.1080/14729679.2011.611283

Collins, L., and Collins, D. (2013). Decision-making and risk management in adventure sports coaching. *Quest* 65, 72–82. doi: 10.1080/00336297.2012.727373

Collins, L., and Collins, D. (2015a). Integration of professional judgement and decision-making in high-level adventure sports coaching practice. *J. Sports Sci.* 33, 622–633. doi: 10.1080/02640414.2014.953980

Collins, L., and Collins, D. (2015b). Professional judgement and decision making in the planning process of high level adventure sports coaching practice. *J. Advent. Outdoor Learn.* doi: 10.1080/14729679.2016.1162182

Collins, L., and Collins, D. (2016a). "Challenges in adventure sports coaching," in *Routledge International Handbook of Outdoor Studies*, eds B. Humberstone, H. Price, and K. A. Hendersson (London: Routledge), 445–462.

Collins, L., and Collins, D. (2016b). Professional judgement and decision-making in adventure sports coaching: The role of interaction. *J. Sports Sci.* 34, 1231–1239. doi: 10.1080/02640414.2015.1105379

Crandall, B., and Getchell-Reiter, K. (1993). Critical decision method: a technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Adv. Nurs. Sci.* 16, 42–51. doi: 10.1097/00012272-199309000-00006

Cruickshank, A., Collins, D., and Minten, S. (2014). Driving and sustaining culture change in Olympic sport performance teams: a first exploration and grounded theory. *J. Sport Exerc. Psychol.* 36, 107–120. doi: 10.1123/jsep.2013-0133

Dane, E., Rockmann, K. W., and Pratt, M. G. (2012). When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organ. Behav. Hum. Decis. Process.* 119, 187–194. doi: 10.1016/j.obhdp.2012.07.009

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Hum. Factors* 37, 65–84.

Endsley, M. R. (1995b). Toward a theory of situational awareness in dynamic systems. *Hum. Factors* 37, 32–64.

Eraut, M. (1994). *Developing Professional Knowledge and Competence.* London: Routledge Falmer.

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *Br. J. Educ. Psychol.* 70, 113–136. doi: 10.1348/000709900158001

Gigerenzer, G., and Todd, P. M. (1998). "Fast and frugal heuristics: the adaptive toolbox," in *Simple Heuristics that Make us smart*, eds G. Gigerenzer, P. M. Todd, and The ABC Research Group (Oxford: Oxford University Press), 3–34.

Harteis, C., and Billett, S. (2013). Intuitive expertise: theories and empirical evidence. *Educ. Res. Rev.* 9, 145–157. doi: 10.1016/j.edurev.2013.02.001

Horn, R. E., and Weber, R. P. (2007). *New Tools for Researching Wicked Problems: Mess Mapping and Resolution Mapping Processes.* Available online at: http://www.strategykinetics.com/New_Tools_For_Resolving_Wicked_Problems.pdf

Janelle, C. M., and Hillman, C. H. (2003). "Expert performance in sport: current perspectives and critical issues," in *Expert Performance in Sports: Advances in Research on Sport Expertise*, eds J. L. Starkes and K. A. Ericsson (Champaign, IL: Human Kinetics), 19–47.

Kaslow, N. J., Rubin, N. J., Bebeau, M. J., Leigh, I. W., Lichtenberg, J. W., Nelos, P. D., et al. (2007). Guiding principles and recommendations for the assessment of competence. *Prof. Psychol.* 38, 441–451. doi: 10.1037/0735-7028.38.5.441

Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *J. Appl. Res. Mem. Cogn.* 4, 164–168. doi: 10.1016/j.jarmac.2015.07.001

Lyle, J., and Cushion, C. (2010). *Sports Coaching: Professionalisation and Practice.* Oxford: Churchill Livingstone.

Martindale, A., and Collins, D. (2007). Enhancing the evaluation of effectiveness with professional judgement and decision making. *Sport Psychol.* 21, 458–474.

Martindale, A., and Collins, D. (2012). A professional judgment and decision making case study: Reflection-in-action research. *Sport Psychol.* 26, 500–518.

Myers, D. G. (2002). *Intuition. Its Powers and Perils.* New Haven, CT: Yale University Press.

Myers, I. B., McCaulley, M. H., Quenk, N. L., and Hammer, A. L. (1998). *The MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator, 2nd Edn.* Palo Alto, CA: Consulting Psychologists Press.

Nonaka, I., and Nishiguchi, T. (eds.). (2001). *Knowledge Emergence: Social, Technical, and Evolutionary Dimensions of Knowledge Creation.* New York, NY: Oxford University Press.

Nonaka, I., and Takeuchi, H. (1995). *The Knowledge-Creating Company.* New York, NY: Oxford University Press.

Pacini, R., and Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs and the ratio-bias phenomenon. *J. Pers. Soc. Psychol.* 76, 972–987. doi: 10.1037/0022-3514.76.6.972

Pretz, J. E. (2011). "Types of intuition: inferential and holistic" in *Handbook of Intuition Research*, ed M. Sinclair (Cheltenham: Edward Elgar Publishers), 17–27.

Pretz, J. E., and Folse, V. N. (2011). Nursing experience and preference for intuition in decision making. *J. Clin. Nurs.* 20, 2878–2889. doi: 10.1111/j.1365-2702.2011.03705.x

Raab, M., and Laborde, S. (2011). When to blink and when to think: preference for intuitive decisions results in faster and better tactical choices. *Res. Q. Exerc. Sport* 82, 89–98. doi: 10.1080/02701367.2011.10599725

Richters, F., Schraagen, J.-M., and Heerkens, H. (2015). Assessing the structure of non-routine decision processes in airline operations control. *Ergonomics.* doi: 10.1080/00140139.2015.1076059. [Epub ahead of print].

Ruth-Sahd, L. A., and Tisdell, E. J. (2007). The meaning and use of intuition in novice nurses: a phenomenological study. *Adult Educ. Q.* 57, 115–140. doi: 10.1177/0741713606295755

Sadler-Smith, E. (2010). *The Intuitive Mind.* West Sussex: Wiley.

Sandelowski, M. (1995). Qualitative analysis: what it is and how to begin. *Res. Nurs. Health* 18, 371–375. doi: 10.1002/nur.4770180411

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. (2002). "The affect heuristic," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, eds T. Gilovich, D. Griffin, and D. Kahneman (Cambridge: Cambridge University Press), 397–420.

Tetlock, P. (2006). *Expert Political Judgement: How good is it? How can we know?* New Jersey, NJ: Princetown University Press

Whitehead, A. E., Taylor, J., and Polman, R. (2015). Examination of the suitability of collecting in event cognitive processes using Think Aloud protocol in golf. *Front. Psychol.* 6:1083. doi: 10.3389/fpsyg.2015.01083

# Cue Utilization and Cognitive Load in Novel Task Performance

*Sue Brouwers[1], Mark W. Wiggins[1]\*, William Helton[2], David O'Hare[3] and Barbara Griffin[1]*

[1] *Macquarie University, Sydney, NSW, Australia,* [2] *University of Canterbury, Christchurch, New Zealand,* [3] *University of Otago, Dunedin, New Zealand*

This study was designed to examine whether differences in cue utilization were associated with differences in performance during a novel, simulated rail control task, and whether these differences reflected a reduction in cognitive load. Two experiments were conducted, the first of which involved the completion of a 20-min rail control simulation that required participants to re-route trains that periodically required a diversion. Participants with a greater level of cue utilization recorded a consistently greater response latency, consistent with a strategy that maintained accuracy, but reduced the demands on cognitive resources. In the second experiment, participants completed the rail task, during which a concurrent, secondary task was introduced. The results revealed an interaction, whereby participants with lesser levels of cue utilization recorded an increase in response latency that exceeded the response latency recorded for participants with greater levels of cue utilization. The relative consistency of response latencies for participants with greater levels of cue utilization, across all blocks, despite the imposition of a secondary task, suggested that those participants with greater levels of cue utilization had adopted a strategy that was effectively minimizing the impact of additional sources of cognitive load on their performance.

Keywords: cue utilization, cognitive resources, cognitive load, workload

## INTRODUCTION

Skilled performance across a range of domains of practice is characterized by accurate and rapid responses, often in dynamic and complex situations (Salthouse, 1991; Ericsson and Lehmann, 1996; Beilock et al., 2004). This is attributed to specialized routines or associations that have been established through repeated application across a variety of settings (Klein, 2011). These highly specialized associations, representative of situation-specific relationships between environmental features and events or objects, are often referred to as cues (Brunswik, 1955; Klein et al., 1986; Wiggins, 2014), and their activation and retrieval from long-term memory has the advantage of imposing relatively fewer demands on working memory resources (Norman and Shallice, 1986; Chung and Byrne, 2008; Evans, 2008).

Differences in the rate at which individuals acquire skills have been attributed to various factors, including cognitive style (Cegarra and Hoc, 2006), motivation and self-regulation (Zimmerman, 2002, 2008), cognitive ability and intelligence (Ackerman, 1986, 2007; Ackerman and Beier, 2007), personality (Singer and Janelle, 1999; Simonton, 2008), and a range of general intrinsic abilities (Thompson et al., 1993; Simonton, 2007, 2008). However, in some environments, the acquisition of skilled performance is also characterized by the capacity to rapidly and accurately extract and utilize meaningful information from features in the environment (Abernethy, 1987, 1990; Bellenkes et al., 1997), thereby enabling the discrimination of relevant from less relevant cues (Weiss and Shanteau, 2003).

Evidence to support the utilization of cues in skill acquisition can be drawn from investigations involving fast ball sports, in which skilled performers anticipate the trajectory of a target by restricting their attention to a limited number of highly predictive features (Müller and Abernethy, 2012; Moore and Müller, 2014). These features include the wrist angle of the bowling arm in cricket (e.g., Müller et al., 2006) and the location of the ball just prior to contact with the racket following a tennis serve (e.g., Jackson and Mogan, 2007).

The rapid identification of a limited number of predictive features has a range of benefits for skill acquisition, including a reduction in the demands on cognitive load and an improvement in the rate of skill acquisition. For example, Perry et al. (2013) were able to demonstrate improvements in performance amongst novice fire fighters by restricting their information acquisition only to those features that were sourced by skilled fire commanders. Although the discrimination between relevant and less relevant features was contrived in this case, it suggests that a general capability to identify a limited number of highly predictive features may explain differences in rates of skill acquisition during unimpeded learning tasks.

Wiggins et al. (2014) demonstrated a relationship between a general capacity for cue utilization and skill acquisition in experiments involving learning to land an aircraft and learning to operate a line-of-sight Unmanned Aerial Vehicle (UAV). Using the situation judgment test EXPERTise (1.0) (Wiggins et al., 2010) to provide a composite assessment of cue-utilization, greater levels of cue utilization were associated with improved accuracy in landing the aircraft following four trials, and with fewer trials to reach criterion in learning to take-off and land a UAV. These improvements in performance occurred in the absence of any formal instruction. However, it was unclear whether these improvements were a consequence of participants' capacity to quickly establish feature-event relationships in the form of cues, and/or whether this capacity reduced the demands on cognitive load, thereby enabling learners to reinforce, revise, or refine the relationships that had been acquired during the initial stages of skill acquisition. The aim of the present study was to investigate, in the context of a low workload, novel task, whether differences in a general capacity for cue utilization are evident in performance, and whether these differences reflect differences in the management of cognitive load.

Where there are multiple courses of action to achieve an outcome, humans will normally select strategies that are associated with the least cognitive effort (Kool et al., 2010). This is referred to as Hull's (1943) law of less work, whereby mental effort is regarded as an aversive stimulus. Therefore, in responding to a novel task, the capacity to identify quickly the strategy of least cognitive effort, while maintaining performance, represents an adaptive approach that conserves cognitive resources.

When exposed to a novel task, participants with a relatively greater capacity for cue utilization would normally be expected to quickly identify key features associated with the performance of a task which, in turn, reduces cognitive load, thereby providing an increased capacity for skill acquisition (Wiggins, 2015). The present study comprised two experiments in the context

of rail control, in which participants were asked to respond to misrouted trains. Importantly, however, participants had seven seconds in which to formulate an assessment, and this represented a key feature that, when identified, would enable participants to minimize the cognitive load imposed by the task.

Consistent with actual rail control, the experimental task was semi-automated, so that it constituted a low workload environment that demanded sustained attention to identify only those trains that required an intervention. Drawing on Resource Theory (Helton et al., 2005; Helton and Warm, 2008), sustained attention to a task is presumed to impose a cognitive demand on information processing, leading to vigilance decrements that include an increase in errors and/or response latency across an extended exposure. Therefore, there was an implicit incentive for participants to adopt a strategy that would reduce cognitive load. In the present study, Experiment 1 examined the relationship between cue utilization and performance on a simulated rail control task over a 20-min period of watch. Experiment 2 involved the imposition of a concurrent secondary task that was intended to, more explicitly, increase cognitive load.

# EXPERIMENT 1

Experiment 1 was designed to examine the relationship between a composite measure of cue utilization, and performance on a simulated rail-monitoring task that required participants to correctly reroute trains that were periodically misrouted. Trains traveled at a consistent and relatively slow rate, and only trains on incorrect routes required a response.

The simulated rail task was designed to incorporate specific elements of ecological validity, including the requirement to monitor multiple rail lines simultaneously, the requirement to intervene periodically, and the requirement to intervene within a specified period of time (Lenior, 1993; Neerincx and de Greef, 1998; Ho et al., 2002; Farrington-Darby et al., 2006). Aside from the adjustment of train routes, which is a fundamental task performed by real-world rail controllers (Neerincx and de Greef, 1998), the movement of trains to and from different directions was also captured in the simulation interface. To account for the demands of experimental control, higher level features of real railway control systems such as the connection of track elements to a network (Berkenkötter and Hannemann, 2006) and the determination/communication of critical incidents (Farrington-Darby et al., 2006) were not incorporated in the simulation task. Given the requirement for sustained attention, the rail-monitoring task continued over a 20-min period of watch. A 20-min period of watch was selected because previous research has found evidence for an observable vigilance decrement within that period of time (Temple et al., 2000; Rose et al., 2002; Helton et al., 2005; Small et al., 2014).

Based on the proposition that a propensity for cue acquisition enables the rapid identification of feature-event relationships, the performance of those participants with relatively greater levels of cue acquisition would, over a consistent period of

exposure to a novel task, be impacted to a relatively lesser extent by the imposition of cognitive load. Since sustained attention is associated with increases in cognitive load (Helton et al., 2005; Helton and Warm, 2008), it was anticipated that, while all participants would experience a vigilance decrement during the latter part of the vigil, participants with greater levels of cue utilization would experience the least increases in response latency coincident with the increase in cognitive load. Specifically, it was hypothesized that: (a) a main effect would be evident for response latency, in which all participants would experience an increase in response latency during the latter stages of the vigil, and (b) that an interaction would be evident, wherein participants with lesser levels of cue utilization would record a greater increase in mean response latency between the first and last 5-min blocks for accurate responses to misrouted trains, in comparison to participants with greater levels of cue utilization.

## Method
### Participants

A total of 58 first and second year university students (41 females and 17 males) were recruited for the study, each of whom received course credit in return for their participation. Participants ranged in age from 18 to 22 years ($M = 19.26$, $SD = 1.35$). The inclusion criteria comprised existing motor vehicle drivers who had not been exposed to train control operations, and who were aged between 18 and 22 years. Utilizing a cohort of 18 to 22 year old drivers enabled comparative assessments of cue utilization, controlling to a limited extent, exposure to driving.

### Instruments

Participants were asked to indicate their age, gender, months of driving experience, daily driving frequency, and their experience in rail control. Cue utilization was assessed using the Expert Skills Evaluation (EXPERTise 1.0) (Wiggins et al., 2010) situation judgment test.

#### EXPERTise 1.0

EXPERTise 1.0 consists of experimental tasks that have been individually and collectively associated with differences in performance at an operational level (Loveday et al., 2013a,b,c, 2014). Consistent with the notion that there are individual differences in populations for cue utilization, the driving version of EXPERTise was selected, as it assesses the acquisition of cues in a specific cohort and at a specific point in time, and it is a context with which participants would be familiar (Wiggins et al., 2014). Tasks in the EXPERTise driving battery include a paired association task, a feature discrimination task, a feature identification task and an information acquisition task.

In the *Paired Association* task, participants are presented with two feature-event/object terms. Over a total of 30 trials, each two terms are displayed, adjacent to one another for 1500 milliseconds. After each pair is displayed, participants indicate the extent to which the two terms are related on a 6-point Likert scale (from 1 = "Extremely unrelated" to 6 = "Extremely

related"). Examples include the related terms 'heavy traffic' (feature) and 'short-cut' (event) and relatively less related terms 'traffic-light' (feature) with 'free-way' (object). Higher levels of cue utilization are associated with a greater variance in the perceived relatedness of terms (Ackerman and Rathburn, 1984; Schvaneveldt et al., 2001; Morrison et al., 2013).

In the *Feature Discrimination* task, participants are presented with a short, written description of a single scenario (i.e., "You are lost in an unfamiliar area. You find yourself in a quiet suburban area, and must find your way to a large shopping center located on a main road. You can see heavier traffic on a main road ahead and high-rise buildings are in the distance..."). Participants are then asked to make a decision based on their typical response in this scenario (i.e., drive in the direction of heavier traffic, or drive toward high-rise housing, and so on). Following their decision, participants are presented with a list of fourteen features and, using a 10-point Likert scale (from 1 = "Not important at all" to 10 = "Extremely important"), are asked to rate these features based on their perceived relevance to his/her decision. Greater levels of cue utilization are associated with higher variances within the feature-relevance ratings (Weiss and Shanteau, 2003; Pauley et al., 2009).

The *Feature Identification* task involves the extraction of key information from an array or scene. Participants are presented with a familiar driving scene (i.e., an image of a road as viewed from the driver's seat of a car) and are directed to identify a road hazard as quickly as possible (i.e., a ball positioned in the road ahead). The position of the ball changes over trials. A lower mean reaction time is associated with greater levels of cue utilization (Schyns, 1998; Schriver et al., 2008; Loveday et al., 2014).

Finally, the *Information Acquisition* task presents participants with a way-finding scenario that requires a choice between three different driving routes. Accompanying the scenario instructions is a drop-down menu with 24 options (feature-cues), which are category-labeled (e.g., 'distance', 'weather conditions') and upon selection, provide participants with information pertaining to the distance, tolls, road works, weather conditions, traffic congestion, speed limit, and the number of lanes for each route. Participants are given one minute to select information prior to making a response. This task assesses the capacity to acquire feature cues from the environment in a prioritized and non-linear pattern (Wiggins and O'Hare, 1995; Wiggins et al., 2002). Individuals with lesser levels of cue utilization are more likely to select information in the sequence in which it is presented (e.g., from left to right as they appear on the display screen). Greater levels of cue utilization are associated with a relatively lower ratio of pairs of information screens accessed in the sequence in which they are presented, against the total frequency of pairs of information screens selected.

The criterion validity of EXPERTise (1.0) has been established in a number of different domains in which typologies formed on the basis of EXPERTise performance differentiated workplace-related performance (Loveday et al., 2013a,b,c). The test–retest reliability ($\kappa = 0.59$, $p < 0.05$) has been demonstrated with power control operators at six-monthly intervals (Loveday et al., 2013a). In the present study, restricting the age of participants (18–22 years) controlled for exposure to driving experience.

This ensured that any differences in cue utilization would be unlikely to result from differences in driving experience. Overall, participants had accumulated a mean of 39 months of driving experience ($SD = 15.82$ months).

### Rail control task

A simulated train control task was used as a novel, low workload context for the present study. In this task, a computer screen depicts a simulated, simplified train control display (see **Figure 1**).

Within the train task display, four long, horizontal green lines represent railway tracks (See **Figure 1**). Each track incorporates an intersection (depicted by white portions on the track), which is controlled by an interlocking switch labeled, "Change". This switch is depicted by a small circle icon, located above each track. If a user selects the "Change" icon, (with a computer mouse), any train traveling on the connected track will be diverted onto the intersecting line.

A train is depicted by a red horizontal bar that appears at one end of a train line, and travels across the display. Each train has a three-digit number assigned as either odd or even (e.g., 888, 333). Each train line and its associated branch line also have an assigned label: Odd or Even. As the train appears onto the screen, a green line depicts the programmed route of the train. The participant's task is to ensure that trains run along the correct train lines (even-numbered trains run along even lines and odd-numbered trains along odd lines). Periodically, programmed routes will appear that are inconsistent with the train's number so that, for example, an even numbered train is programmed to take a route that is labeled 'odd'. To correct the programmed route of the train, participants must select the "Change" icon which will re-route the train.

Once a train appears on the computer screen, participants have seven seconds in which to decide whether or not to reroute a train. All trains travel at the same speed and trains appear within 5–30 s of each other. Therefore, the screen may display a static image of train lines (without any trains) for up to 30 s before another train appears. A total of 67 trains appear on the four rail lines over the course of 20-min, half of which are not required to be re-routed. Data recorded from this task included response latency (in milliseconds, from the initial appearance of a train, to the selection of the "Change" icon) and the accuracy of responses (whether trains were diverted when required).

### Cognitive ability

The Raven's Standard Progressive Matrices cognitive test (SPM; Raven et al., 1998, 2000) was included as a measure of cognitive ability. The SPM broadly assesses general problem solving ability or fluid intelligence by measuring the capacity to recognize and process patterns of spatial information (Raven et al., 2000; Kaplan and Saccuzzo, 2008). Cognitive ability encompasses constructs that include processing speed and working memory capacity (Conway et al., 2002) that can influence performance in attention-demanding tasks (Kane and Engle, 2003). In the present study, the SPM was included as a means of establishing whether cognitive ability was related to performance scores in the rail task. The SPM short version (10-min timed) was used (see Caffarra et al., 2003; Austin, 2005; Moutafi et al., 2006; Jaeggi et al., 2011). Cognitive ability scores reflected the total number of correct SPM responses.

### The group embedded figures test

The Group Embedded Figures Test (GEFT: Witkin et al., 1971, 2002; Oltman et al., 2003) is a perceptual test that assesses an individual's field dependence-independence. According to Witkin (1976), Field Independence–Dependence is a cognitive style that represents the extent to which an individual can overcome the influence of irrelevant background elements when
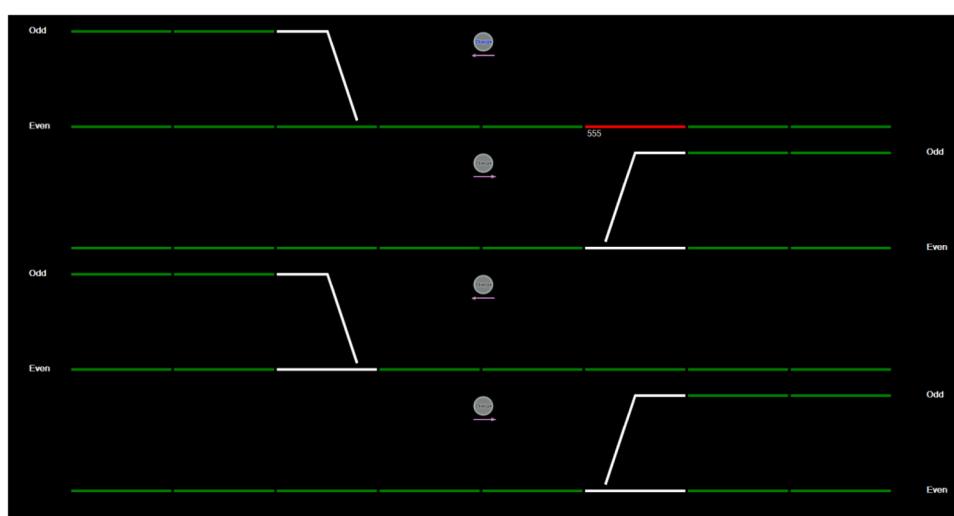


**FIGURE 1 | The simulated train-control display as viewed by participants.** The four long, horizontal green lines represent railway tracks. The white portions on each track are the intersection lines, which are controlled by an interlocking switch labeled, "Change". This switch is depicted by a small circle icon, located above each track. If a participant selects the "Change" icon, any train traveling on the track beneath it, will be diverted onto the intersecting line.

attending to a task. Individuals who exhibit higher levels of field independence more easily overcome background elements in formulating judgments. The GEFT requires the test taker to identify and trace simple forms (i.e., shapes) that are embedded within more complex forms. The Embedded Figures Test has been linked to the capacity to perceive hazards, recognize faults and formulate mental representations of problems (Vessey and Galletta, 1991; Elander et al., 1993; Leach and Morris, 1998). The GEFT was included in the present study to ascertain whether rail task responses were related to cognitive style. Test–retest reliability coefficients for the GEFT range from 0.79 to 0.92 over multiple time intervals of up to 3 years (Kepner and Neimark, 1984; Witkin et al., 2002).

## Procedure

Following approval of the study by the Macquarie University Human Research Ethics Committee, participants were recruited and tested individually in 90-min sessions. After completing an on-line demographic questionnaire, a computer prompt directed the participants through the four EXPERTise tasks. Standardized instructions for the rail task were then provided verbally. This included the verbal instruction, "the aim of this task is to ensure that each train is on its correct track". No information or direction was provided in relation to the speed or pace of the task (i.e., participants were not told that they had several seconds of decision-time available or that they could or should respond in either an immediate or delayed manner). After a 5-min trial to orient the participants to the task, the, 20-min experimental trial commenced. Participants then completed paper-and-pencil versions of the SPM and GEFT. Instructions for these tests were provided to participants verbally and through written directions, according to the test instruction manuals.

## Results

### Preliminary Analysis

#### Rail task performance scores

Response latency for correct responses in the rail task comprised the primary dependent variable. Latencies were calculated from the initial appearance of a train to the selection of the 'change' icon where appropriate. Errors occurred when a train was re-routed from its correct path (a false alarm) or was not re-routed when required (a miss). The number of errors made by participants ranged from zero to five, with a median of one, and resulted in a floor effect, with 64% of the entire sample recording either zero or a single error during exposure to the 67 trains. A Spearman's rank-ordered, non-parametric correlation between the number of errors committed in the rail task and mean response latencies was not statistically significant. The relationship between error frequency and interval, examined using a chi-square test of independence, failed to reveal any statistically significant variation in the distribution of errors across the four time intervals, $\chi^2(3, 58) = 5.026$, $p = 0.17$. Taken together, these results suggest that a speed-accuracy trade-off was not necessary to undertake the task successfully.

Since the task was 20-min in duration, the mean response latencies (for correct responses) were calculated across four, 5-min intervals, and these four variables comprised the dependent variables in subsequent analyses. Nineteen trains appeared within the first block, nine of which required re-routing. In the second block, 16 trains appeared, eight of which required re-routing. In the third block, 15 trains appeared, seven of which required re-routing, and in the final time block 17 trains appeared, of which nine required re-routing.

### Cognitive ability and cognitive style

Scores on the SPM were normally distributed and not significantly correlated with mean response latencies for any of the four blocks of trials ($-0.04 \leq r \leq -0.15$, $p > 0.05$). As GEFT (cognitive style) scores were negatively skewed, a square root transformation with reflection was applied to normalize the data. Subsequent Pearson's correlations failed to reveal any statistically significant associations between GEFT scores and mean response latencies across any of the four blocks of trials ($-0.03 \leq r \leq 0.22$, $p > 0.05$).

### Cue utilization typologies

Prior to analysis, it was necessary to identify the cue utilization typologies that corresponded to relatively greater or lesser levels of cue utilization (Loveday et al., 2013a,b; Wiggins et al., 2014). Consistent with the standard approach to EXPERTise data, $z$ scores were calculated for each task, with those corresponding to the Information Acquisition and Feature Identification tasks reversed so that for all four tasks, higher $z$ scores represented greater levels of cue utilization. A cluster analysis identified two groups with centroids corresponding to higher variance in the Paired Association and Feature Discrimination tasks, lower response latency in the Feature Identification task (reversed $z$ score), and a lower ratio of sequential selections in the Information Acquisition task (reversed $z$ score). The cluster analysis classified 34 participants in the lesser cue utilization typology and 24 participants in the greater cue utilization typology (**Table 1**).

### Driving experience and cue utilization

To examine whether differences in cue utilization resulted from differences in participants' length of driving experience, a one-way Analysis of Variance (ANOVA) was conducted using EXPERTise cluster as the independent variable, and months of driving experience as the dependent variable. The length of driving experience reported by participants in the lesser cue utilization cluster ($M = 38.24$, $SD = 12.69$) did not differ significantly from those participants with greater levels of cue utilization ($M = 39.50$, $SD = 19.70$), $F(1,57) = 0.088$, $p = 0.77$,

TABLE 1 | Cluster centroids for the EXPERTise task scores.

|  | Cluster 1 (*n* = 34) | Cluster 2 (*n* = 24) |
| --- | --- | --- |
| Paired association | −0.60 | 0.86 |
| Feature discrimination | −0.52 | 0.74 |
| Feature identification | −0.12 | 0.17 |
| Information acquisition | −0.40 | 0.57 |

suggesting that assessments of cue utilization were not related to driving exposure.

## Cue Utilization and Rail Task Performance

The primary aim of the present study was to establish whether differences existed between levels of cue utilization (cue typologies) and response latency across these four rail-control task blocks (a time block × cue typology interaction). A 2 × 4 mixed ANOVA, comprising two levels of cue utilization (greater and lesser) as a between-groups factor and four blocks of trials as a within-groups variable failed to reveal a statistically significant interaction between the variables, $F(2.62, 146.56) = 1.09$, $p = 0.349$, $\eta_p^2 = 0.019$. This suggests that the changes evident in the mean response latency over trials occurred at similar rates, irrespective of cue utilization typology.

Despite the fact that an interaction was not evident between cue utilization typology and blocks of trials, main effects were, nevertheless, evident for cue utilization typology, $F(1, 56) = 20.36$, $p < 0.001$, $\eta^2 = 0.267$ and for blocks of trials, $F(2.60, 147.89) = 7.37$, $p = 0.001$, $\eta^2 = 0.114$. Inspection of the mean response latencies (**Figure 2**) indicated that participants with a greater level of cue utilization recorded a *slower* mean response latency ($M = 2079.70$, $SD = 395.67$, $SE = 80.77$) across the four blocks of the rail-control task, in comparison to participants with a lesser level of cue utilization ($M = 1527.36$, $SD = 498.59$, $SE = 85.51$). Since there were no differences in the accuracy of the two groups, it suggests that participants with greater levels of cue utilization either withdrew cognitive resources to reduce the demand on cognitive load, or alternatively, invested cognitive resources to maintain accuracy.

*Post-hoc* analysis of the mean response latencies for blocks of trials indicated that mean response latencies in the first block of trials ($M = 1595.51$, $SD = 558.33$, $SE = 73.31$) were significantly lower than the fourth block ($M = 1921.37$, $SD = 687.93$, $SE = 90.33$), $t(57) = -3.87$, $p < 0.001$. This increase in mean response latency over time, despite no changes in task requirements, is consistent with the vigilance decrement.



**FIGURE 2 | Rail task response latencies by cue utilization typology and block number for Experiment 1.** Error bars represent ±1 *SE*.

## Discussion

This study was designed to examine whether, in response to a novel, short vigilance task, participants with a greater capacity for cue acquisition would adopt a strategy that would reduce the demands on cognitive resources. It was hypothesized that a strategy of least cognitive effort would be evident in an interaction that would emerge as the train control task progressed. On the basis of the Resource Theory explanation of the vigilance decrement, it was assumed that the increase in cognitive load that is associated with an extended period of watch would differentially affect those participants with lesser levels of cue utilization. Although a main effect was evident with progressive increases in response latency across blocks of trials, consistent with the hypothesized vigilance decrement, no statistically significant interaction occurred.

A main effect of cue utilization was also evident in which participants with a greater level of cue utilization showed increased response latencies in response to the diversion of trains. These mean response latencies were not associated with either cognitive ability (SPM scores) nor cognitive style (GEFT scores). However, it was unclear whether this response resulted in a reduction in cognitive load. Since there were no differences in the accuracy of responses amongst the two groups, the results suggest that participants with greater levels of cue utilization recognized that time was available in which to initiate a response to reroute misrouted trains, and adopted a strategy of least cognitive effort.

Although greater levels of cue utilization are normally associated with a reduction in response latency, this is not always the case. For example, in self-paced, targeting tasks such as rifle shooting and basketball (free throwing), superior shot accuracy is associated with longer quiet eye periods (the final fixation on the target prior to the initiation of movement) (Vickers, 1996; Vickers and Williams, 2007). As a result, skilled players tend to take more time to execute shots than lesser skilled players (Williams et al., 2002; Vickers, 2007). This suggests that the advantage afforded by greater levels of cue utilization lies in the capacity to recognize the need to adapt to different task demands. In the present study, there was no loss of performance associated with the increased response latency and it may have constituted a strategy of least cognitive effort which enabled the maintenance of performance despite the increase in cognitive demands.

There are at least two explanations for the lack of an interaction between levels of cue utilization and blocks of trials, the first of which relates to the hypothesized reduction in cognitive load. In particular, the self-pacing of one's actions and responses within a task or job has been identified as a workload management strategy that effectively increases task control and reduces cognitive demands and anxiety (Johansson, 1981; Salvendy and Smith, 1981; Scerbo et al., 1993). However, it may be the case that the workload demands in the present study were insufficient to draw on the cognitive resources that would have been necessary to differentiate participants with greater or lesser levels of cue utilization.

An alternative explanation for the lack of an interaction relates to a potential investment of cognitive resources amongst participants with greater levels of cue utilization. Specifically,
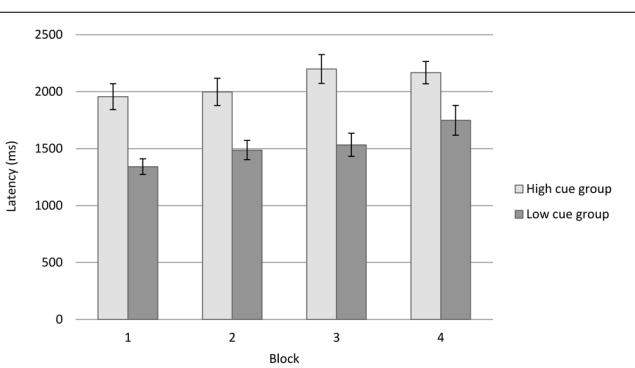
it might be argued that greater attention to the task, although overcompensating for the resources necessary to maintain accuracy, resulted in the increase in response latency. Experiment 2 was designed to differentiate the two explanations through the imposition of a secondary task that explicitly increased the cognitive demands during the rail control simulation.

## EXPERIMENT 2

Consistent with Experiment 1, participants in Experiment 2 completed the EXPERTise 1.0 situation judgment test and the 20-min simulated rail-control task. However, in addition to monitoring the rail display and re-routing trains as necessary, participants in Experiment 2 were asked to complete a secondary task during the final two blocks (10-min) of trials that comprised the monitoring task. This secondary task was designed to impose an explicit cognitive load, and required individuals to note the assigned number of each train (i.e., 888), together with the time at which it appeared (i.e., 2.07 PM).

Assuming that the advantage afforded by greater levels of cue utilization during the performance of a novel task is a reduction in cognitive load, it was anticipated that the imposition of a secondary task would impact the performance of participants with greater or lesser levels of cue utilization differently and at different stages of the task. It was hypothesized that an interaction would be evident in which participants with lesser levels of cue utilization would record an increase in response latency, while no effect would be evident for participants with greater levels of cue utilization.

## Method
### Participants
Fifty-nine university students (15 males and 44 females) aged between 18 and 22 years ($M = 18.81$, $SD = 1.06$) participated in the study and received course credit for their participation. As in Experiment 1, individuals were excluded if they were not existing drivers, had acquired experience in the context of rail control, or were outside of the 18–22 year-old inclusion range. Participants in Experiment 1 of the study were also excluded from participating in Experiment 2.

### Instruments
#### EXPERTise
The same four driving EXPERTise tasks (Wiggins et al., 2010) utilized in Experiment 1, were included as a composite measure of driving-related cue utilization across four cue-based problem solving and processing dimensions. An additional *Feature Identification* task was included, which exposed participants to a series of 18 different road images (photographs), each displayed for 500 ms, and required participants to estimate the speed limit of each road from four multiple-choice options (50–60, 70–80, 90–100 or 110+ km/hr). Designed to assess the capacity to rapidly extract key information from a driving-related scene and form an accurate judgment, a greater number of accurate judgments in this task was expected to reflect greater levels of cue utilization.

#### Rail control task
Participants in Experiment 2 completed the simulated train control task that was used in Experiment 1. However, in Experiment 2, participants completed the final two, 5-min blocks in conjunction with a secondary task.

#### Secondary task
A manipulation check was undertaken with five volunteers to ensure that the secondary task reduced the decision-time afforded to participants in the rail task, but did not induce an extremely low or an impossibly high level of workload such that the accuracy of responses would be impacted. The secondary task required participants to write down the train number and the time at which each train appeared on the screen. Following a 5-min period of familiarization, three volunteers completed the first half of the rail task (10-min) with the inclusion of the secondary task, while two volunteers completed the second half of the rail task (10-min) with the inclusion of the secondary task. Trials were counterbalanced to control for sequencing effects, such as fatigue, that were unrelated to the secondary task. The manipulation check revealed no errors in the secondary task (all trains were correctly logged), while response latency was greater for the dual task condition ($M = 3063$ ms) compared to the vigil-only condition ($M = 2691$ ms) suggesting that the secondary task increased the workload to an adequate but not extreme degree.

#### Subjective workload
Subjective workload was measured by the NASA Task Load Index (NASA-TLX: Hart and Staveland, 1988), a widely used and validated multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental demands, physical demands, temporal demands, performance, effort, and frustration (Hart and Staveland, 1988; Xiao et al., 2005) on a scale of 1–100. Participants completed the NASA-TLX following the single rail-task condition (Blocks 1 and 2) and again following the secondary task condition (Blocks 3 and 4).

### Procedure
As in Experiment 1, participants were tested individually and completed the study in sessions of 90 min. Following the completion of a demographic questionnaire, participants undertook the EXPERTise tasks and a 5-min practice trial to orient participants to the rail task. Prior to the rail control task, instructions were provided to participants in relation to the distractor task and they were given the paper-based secondary-task sheet. Once participants indicated that the instructions were understood, the simulated rail control task commenced. After 10 minutes, the rail task was paused by the researcher and participants completed the NASA-TLX. The rail task then recommenced, and for the remaining ten minutes of the task, participants diverted trains and completed the secondary-task sheet concurrently. Following the completion of the rail task, participants again completed the NASA-TLX.

# Results

## Cue Utilization Typologies

Consistent with Experiment 1, a cluster analysis was undertaken using aggregated EXPERTise z scores for all five tasks to identify the cue utilization typologies that corresponded with relatively greater and lesser levels of cue utilization. Two groups were identified with centroids corresponding to higher variance in the Paired Association and Feature Discrimination tasks, lower response latency in the Feature Identification tasks (reversed z scores), and lower ratio of sequential selections in the Information Acquisition task (reversed z score). In this case, the cluster analysis (**Table 2**) classified 22 participants in the lower cue utilization typology (cluster 1) and 33 participants in the higher cue utilization typology (cluster 2).

## Driving Experience and Cue Utilization

Consistent with Experiment 1, the duration of driving experience (months) reported by participants in the lesser cue utilization cluster ($M = 29.73$, $SD = 13.06$) did not differ significantly from those participants who were classified in the greater cue utilization cluster ($M = 29.57$, $SD = 13.60$), $F_{(1,50)} = 0.002$, $p = 0.97$. This suggests that differences in cue utilization did not result from differences in participants' driving experience.

## Rail Task Performance

Consistent with the results in Experiment 1, a floor effect was evident for the frequency of errors during the rail control task (range = 0–4, $Mdn = 1$) with 68% of participants committing either zero or a single error during exposure to 67 trains. A Chi-square test of independence indicated there were no significant differences in the distribution of errors across the four time intervals, $\chi^2_{(3,59)} = 5.78$, $p = 0.123$. The frequency of errors committed was unrelated to response latencies (Spearman's non-parametric, $0.18 \leq r \leq 0.26$, $p > 0.05$).

## Cue Utilization and Rail Task Latencies

To investigate whether the imposition of the secondary task had a greater impact on participants with lesser levels of cue utilization compared to those participants with greater levels, a 2 × 4 mixed repeated ANOVA was undertaken, including the two levels of cue utilization (greater, lesser) as a between-groups variable and the four blocks of trials as a within groups variable. Consistent with the hypothesis, an interaction was evident between cue utilization and block trials, $F_{(1.80,90.21)} = 10.81$, $p < 0.001$, $\eta_p^2 = 0.178$ (Greenhouse–Geisser correction), in which the mean response latency for participants increased with lesser levels of cue utilization, while the mean response latency for participants with
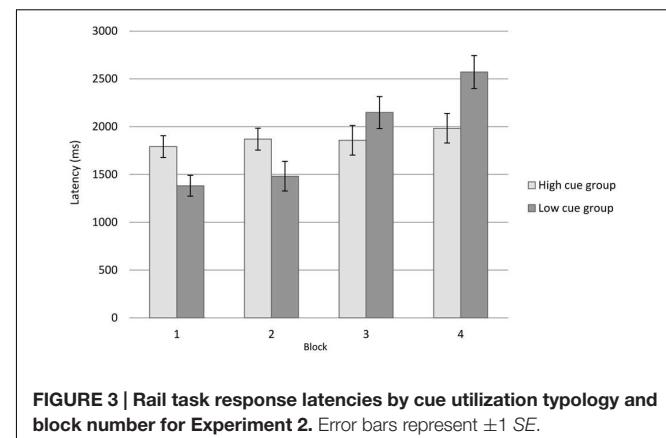


**FIGURE 3 | Rail task response latencies by cue utilization typology and block number for Experiment 2.** Error bars represent ±1 SE.

greater levels of cue utilization remained relatively consistent (**Figure 3**). This suggests that the imposition of the secondary task had a greater impact on participants with lesser levels of cue utilization in comparison to participants with greater levels of cue utilization.

A main effect was evident for blocks of trials, $F_{(1.65,95.72)} = 12.11$, $p < 0.001$, $\eta_p^2 = 0.173$. *Post hoc* analysis of the mean response latencies for blocks of trials indicated that the mean response latencies in the first block of trials ($M = 1608.56$, $SD = 594.66$, $SE = 77.42$) were significantly lower than in the final block of trials ($M = 2226.61$, $SD = 851.81$, $SE = 110.90$), $t_{(58)} = -4.51$, $p < 0.00$. The main effect of cue utilization was not statistically significant, $F_{(1,50)} = 0.17$, $p = 0.90$.

As is evident from **Figure 3**, the pattern of response latencies following the imposition of the secondary task differed on the basis of levels of cue utilization. This suggests that the relative impact of the secondary task was greatest for participants with lesser levels of cue utilization than was the case for participants with greater levels of cue utilization.

## Cue Utilization and Mental Workload Perceptions

To investigate whether the imposition of the secondary task impacted participants' perceptions of mental workload, a 2 × 2 mixed repeated ANOVA was undertaken, with cue utilization level (greater and lesser) as the between-groups factor and TLX scores (single-condition and dual-condition) as the within-groups variable. The results revealed a statistically significant main effect for perceptions of mental workload, $F_{(1,50)} = 85.33$, $p < 0.001$, $\eta_p^2 = 0.631$, in which participants perceived the task workload in the dual condition as significantly greater ($M = 26.83$, $SD = 1.90$), than during the single task condition ($M = 14.78$, $SD = 1.40$), $t_{(58)} = -9.22$, $p < 0.001$. There was no main effect for cue utilization, $F_{(1,50)} = 0.58$, $p = 0.449$.

Consistent with the results pertaining to response latency, a statistically significant interaction was evident between perceptions of mental workload and cue utilization, $F_{(1,50)} = 8.00$, $p = 0.007$, $\eta_p^2 = 0.138$. As is evident from **Figure 4**, the pattern of perceived mental workload (as measured by the NASA-TLX) following the imposition of the secondary task differed on the basis of levels of cue utilization. Specifically,

**TABLE 2 | Cluster centroids for the EXPERTise task scores.**

|  | Cluster 1 (n = 22) | Cluster 2 (n = 33) |
| --- | --- | --- |
| Paired association | −0.83 | 0.56 |
| Feature discrimination | −0.84 | 0.53 |
| Feature identification | −0.30 | 0.21 |
| Feature identification II | −0.45 | 0.33 |
| Information acquisition | −0.18 | 0.18 |
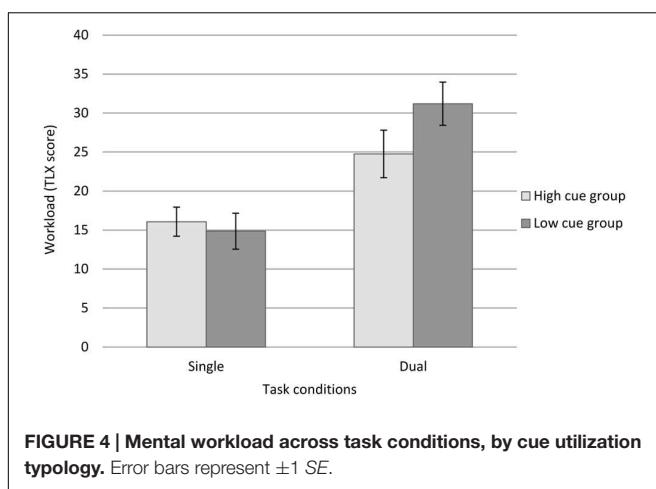
**FIGURE 4 | Mental workload across task conditions, by cue utilization typology.** Error bars represent ±1 SE.

the perceived impact of the secondary task was greatest for participants with lesser levels of cue utilization.

## Discussion

The introduction of the secondary task part-way during the 20-min period of rail control was designed to impose an explicit cognitive demand on the performance of participants. It was reasoned that if participants with greater levels of cue utilization had adopted a strategy that effectively reduced the demands on cognitive resources, then an interaction should be evident following the introduction of the secondary task during the final two, 5-min blocks of the 20-min trial. Specifically, it was hypothesized that participants with lesser levels of cue utilization would record an increase in response latency, while only a minimal effect would be evident for participants with greater levels of cue utilization. Consistent with the hypothesis, mean response latencies for participants with lesser levels of cue utilization increased following the introduction of the secondary task and continued to increase as the task progressed, while the mean response latencies for participants with greater levels of cue utilization remained consistent with the vigilance decrement that was evident in Experiment 1. This effect occurred independent of driving experience but was reflected in perceptions of mental workload.

## GENERAL DISCUSSION

In response to a novel task, the rapid development of associational cues in memory is one means by which the cognitive demands of a task can be minimized (Norman and Shallice, 1986; Chung and Byrne, 2008; Evans, 2008). The aim of the research presented in this paper was to examine whether differences in cue utilization were associated with differences in performance during a novel, simulated rail control task, and whether these differences in performance reflected a reduction in cognitive load. On the assumption that cognitive load increases with sustained attention to a task (Helton et al., 2005; Helton and Warm, 2008), it was anticipated that individuals with relatively greater levels of cue utilization would be relatively less impacted by the sustained

attentional demands imposed by a simulated rail-control task in which participants were asked to identify and correct the path of trains that had periodically been misrouted.

Two experiments were conducted with motor vehicle drivers aged between 18 and 22 years who undertook an assessment of cue utilization using the driving battery of EXPERTise 1.0. In Experiment 1, participants who were identified a priori with a relatively greater level of cue utilization on the basis of their scores on EXPERTise 1.0, recorded a mean response latency greater than that recorded by participants with relatively lesser levels of cue utilization. The effect remained consistent across the four blocks of 5-min trials within the rail-control task. Importantly, there were no differences in accuracy and, in fact, a floor effect was evident in relation to errors.

A vigilance decrement was evident in the increases in response latency recorded across blocks of trials, irrespective of participants' level of cue utilization. This suggests that, although an increase in cognitive load may have been associated with sustained attention to the task, the level was insufficient to differentiate the performance of participants on the basis of their cue utilization. Consequently, Experiment 2 adopted a similar methodology but included a secondary task to invoke an explicit cognitive load part-way through the simulated rail control task.

The performance of participants in Experiment 2 during the initial two blocks of trials appeared consistent with the results from Experiment 1, whereby the response latency recorded was higher for participants with greater levels of cue utilization. However, once the secondary task was initiated, the response latency of participants with lesser levels of cue utilization increased, while the response latency amongst participants with greater levels of cue utilization remained relatively consistent. This suggests that the relative impact of the secondary task was greater for participants with lesser levels of cue utilization than it was for participants with greater levels of cue utilization.

The relative consistency of response latencies recorded for participants with higher levels of cue utilization across all blocks despite the imposition of a secondary task, suggests that they had adopted a strategy that reduced the demands on cognitive load. Until the introduction of a secondary task, the mean response latency for participants with greater levels of cue utilization was consistently greater than the mean response latency recorded by participants with lesser levels of cue utilization. Therefore, it might be concluded that participants were adopting a strategy of self-pacing, which effectively increased task control and reduced cognitive demands (Johansson, 1981; Salvendy and Smith, 1981; Scerbo et al., 1993). As a decision to re-route trains in the rail simulation task could be initiated up to seven seconds from the appearance of a train, those participants with greater levels of cue utilization appear to have recognized this opportunity and utilized the additional time, without sacrificing accuracy.

In contrast, the pattern of results for those participants with lesser levels of cue utilization, suggests that, until the imposition of the secondary task, these participants may have been responding rapidly and reactively, rather than in a manner consistent with the strategic conservation of resources to manage workload (Hollnagel, 2002; Hollnagel and Woods, 2005;

Loft et al., 2007). Their rapid increase in mean response latencies subsequent to the imposition of the secondary task suggested that their reactive responses were unable to be sustained with the increasing level of workload.

It is noteworthy, however, that those participants with lesser levels of cue utilization maintained consistent (and low) levels of error rates throughout the rail task, and this occurred despite the increased workload imposed by the secondary task. Therefore, it is also possible that those participants with lesser levels of cue utilization may have adopted a strategy that increasingly sacrificed speed for accuracy. Given that the workload of the task imposed demands that did not impact accuracy, it is likely that a further increase in cognitive demands would, despite efforts to minimize effort, exhaust the information processing resources of those participants with lesser levels of cue utilization and result in a deterioration in accuracy. To explore if this is the case, future research may consider increasing the level of cognitive demand by either extending the duration of the vigil (e.g., Freeman et al., 2004; Nelson et al., 2014) or increasing the demands of the task (Matthews and Davies, 1998; Smit et al., 2004) to a point where accuracy is impeded (Smit et al., 2004).

Overall, the results of both experiments provide support for the assertion that a relatively greater capacity for cue utilization is associated with an increased capacity to cope with the demands of a novel task. Throughout both experiments, several control measures were utilized to ensure that performance differences between individuals with lesser and greater levels of cue utilization were not due to cognitive ability nor cognitive style. These variables were not related to response latencies. Consistent with previous research (Smeeton et al., 2004; Müller and Abernethy, 2012; Moore and Müller, 2014; Wiggins et al., 2014), our results suggest that a propensity to identify critical cues and rapidly establish feature-event relationships may provide an opportunity to reduce cognitive demands, thereby enabling the acquisition of new features and/or the opportunity to revise or refine existing features.

In practice, implications that arise from the present study present tangible opportunities in the context of selection and training. The ability to identify the levels of cue utilization may provide the basis to differentiate job applicants that are more or less likely to acquire skills in the absence of a dedicated training regime. The outcomes might also be applied to identify employees who are most in need of a training intervention, particularly in the context of the identification of key features that might enable a reduction in cognitive load and the subsequent acquisition and revision of feature-event relationships in the form of cues (Wulf et al., 2000; Lagnado et al., 2006).

What remains to be established is the extent to which the association between cue utilization and performance evident in the present research can be generalized. For example, driving and rail control both involve visual perception and spatial skills. The driving version of EXPERTise may be less capable of differentiating performance beyond this context. It is also noteworthy that while the results of this study suggest that participants with a greater capacity for cue utilization adopted a strategy that minimized the impact of additional cognitive load on their performance, the precise nature of that strategy (which may pertain to the utilization of available time to self-pace) has yet to be investigated and explicated.

## CONCLUSION

The present study was designed to examine whether differences in cue utilization were associated with differences in performance during a novel, simulated rail control task, and whether these differences in performance reflected a reduction in cognitive load. The results of two experiments suggested that levels of cue utilization were associated with differences in response latencies throughout the simulated rail task, and that individuals with a greater level of cue utilization were able to adopt a strategy that effectively reduced cognitive load without sacrificing accuracy.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENT

## REFERENCES

Abernethy, B. (1987). Anticipation in sport: a review. *Phys. Educ. Rev.* 10, 5–16.

Abernethy, B. (1990). Anticipation in squash: differences in advance cue utilization between expert and novice players. *J. Sports Sci.* 8, 17–34. doi: 10.1080/02640419008732128

Ackerman, B. P., and Rathburn, J. (1984). The effect of recognition experience on cued recall in children and adults. *Child Dev.* 55, 1855–1864. doi: 10.2307/1129932

Ackerman, P. L. (1986). Individual differences in information processing: an investigation of intellectual abilities and task performance during practice. *Intelligence* 10, 101–139. doi: 10.1016/0160-2896(86)90010-3

Ackerman, P. L. (2007). New developments in understanding skilled performance. *Curr. Dir. Psychol. Res.* 16, 235–239. doi: 10.1111/j.1467-8721.2007.00511.x

Ackerman, P. L., and Beier, M. E. (2007). Further explorations of perceptual speed abilities, in the context of assessment methods, cognitive abilities and individual differences during skill acquisition. *J. Exp. Psychol.* 13, 249–272. doi: 10.1037/1076-898X.13.4.249

Austin, E. J. (2005). Emotional intelligence and emotional information processing. *Pers. Individ. Differ.* 39, 403–414. doi: 10.1016/j.paid.2005.01.017

Beilock, S. L., Bertenthal, B. I., McCoy, A. M., and Carr, T. H. (2004). Haste does not always make waste: expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychon. Bull. Rev.* 11, 373–379. doi: 10.3758/BF03196585

Bellenkes, A. H., Wickens, C. D., and Kramer, A. F. (1997). Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviat. Space Environ. Med.* 68, 569–579.

Berkenkötter, K., and Hannemann, U. (2006). "Modeling the railway control domain rigorously with a UML 2.0 profile," in *Computer Safety, Reliability, and Security*, eds F. Saglietti and N. Oster (Berlin: Springer), 398–411.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychol. Rev.* 62, 193–271. doi: 10.1037/h0047470

Caffarra, P., Vezzadini, G., Zonato, F., Copelli, S., and Venneri, A. (2003). A normative study of a shorter version of Raven's progressive matrices 1938. *Neurol. Sci.* 24, 336–339. doi: 10.1007/s10072-003-0185-0

Cegarra, J., and Hoc, J. M. (2006). Cognitive styles as an explanation of experts' individual differences: a case study in computer-assisted troubleshooting diagnosis. *Int. J. Hum. Comput. Stud.* 64, 123–136. doi: 10.1016/j.ijhcs.2005.06.003

Chung, P. H., and Byrne, M. D. (2008). Cue effectiveness in mitigating post completion errors in a routine procedural task. *Int. J. Hum. Comput. Stud.* 66, 217–232. doi: 10.1016/j.ijhcs.2007.09.001

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., and Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence* 30, 163–183. doi: 10.1016/S0160-2896(01)00096-4

Elander, J., West, R., and French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: an examination of methods and findings. *Psychol. Bull.* 113, 279–294. doi: 10.1037/0033-2909.113.2.279

Ericsson, K. A., and Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu. Rev. Psychol.* 47, 273–305. doi: 10.1146/annurev.psych.47.1.273

Evans, J. S. B. (2008). Dual-processes accounts of reasoning. *Annu. Rev. Psychol.* 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629

Farrington-Darby, T., Wilson, J. R., Norris, B. J., and Clarke, T. (2006). A naturalistic study of railway controllers. *Ergonomics* 49, 1370–1394. doi: 10.1080/00140130600613000

Freeman, F. G., Mikulka, P. J., Scerbo, M. W., and Scott, L. (2004). An evaluation of an adaptive automation system using a cognitive vigilance task. *Biol. Psychol.* 67, 283–297. doi: 10.1016/j.biopsycho.2004.01.002

Hart, S. G., and Staveland, L. E. (1988). "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research," in *Human Mental Workload*, eds P. A. Hancock and N. Meshkati (Amsterdam: North-Holland), 139–183.

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., et al. (2005). Signal regularity and the mindlessness model of vigilance. *Br. J. Psychol.* 96, 249–261. doi: 10.1348/000712605X38369

Helton, W. S., and Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychol.* 129, 18–25. doi: 10.1016/j.actpsy.2008.04.002

Ho, T. K., Mao, B. H., Yuan, Z. Z., Liu, H. D., and Fung, Y. F. (2002). Computer simulation and modeling in railway applications. *Comput. Phys. Commun.* 143, 1–10. doi: 10.1016/S0010-4655(01)00410-6

Hollnagel, E. (2002). Time and time again. *Theor. Issues Ergon. Sci.* 3, 143–158. doi: 10.1080/14639220210124111

Hollnagel, E., and Woods, D. D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. Boca Raton, FL: CRC Press.

Hull, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century-Crofts.

Jackson, R. C., and Mogan, P. (2007). Advance visual information, awareness, and anticipation skill. *J. Mot. Behav.* 39, 341–351. doi: 10.3200/JMBR.39.5.341-352

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108

Johansson, G. (1981). "Psychoneuroendocrine correlates of unpaced and paced performance," in *Machine Pacing and Occupational Stress*, eds G. Salvendy and M. J. Smith (London: Taylor & Francis), 277–286.

Kane, M. J., and Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J. Exp. Psychol.* 132, 47–70. doi: 10.1037/0096-3445.132.1.47

Kaplan, R. M., and Saccuzzo, D. P. (2008). *Psychological Testing: Principles, Applications, and Issues*, 7th Edn. Belmont, CA: Wadsworth Thomson Learning.

Kepner, M. D., and Neimark, E. D. (1984). Test–retest reliability and differential patterns of score change on the Group Embedded Figures Test. *J. Pers. Soc. Psychol.* 46, 1405–1413. doi: 10.1037/0022-3514.46.6.1405

Klein, G. (2011). "Intuition and naturalistic decision-making," in *Handbook of Intuition Research*, ed. M. Sinclair (Cheltenham: Edward Elgar Publishing), 69–78.

Klein, G. A., Calderwood, R., and Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. *Hum. Fact. Ergon. Soc. Ann. Meet. Proc.* 30, 576–580. doi: 10.1177/154193128603000616

Kool, W., McGuire, J. T., Rosen, Z. B., and Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *J. Exp. Psychol.* 139, 665–682. doi: 10.1037/a0020198

Lagnado, D. A., Newell, B. R., Kahan, S., and Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *J. Exp. Psychol.* 135, 162–183. doi: 10.1037/0096-3445.135.2.162

Leach, J., and Morris, P. E. (1998). Cognitive factors in the close visual and magnetic particle inspection of welds underwater. *Hum. Fact.* 40, 187–197. doi: 10.1518/001872098779480460

Lenior, T. M. J. (1993). Analyses of cognitive processes in train traffic control. *Ergonomics* 36, 1361–1368. doi: 10.1080/00140139308968005

Loft, S., Sanderson, P., Neal, A., and Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: critical review and broader implications. *Hum. Factors* 49, 376–399. doi: 10.1518/001872007X197017

Loveday, T., Wiggins, M. W., Festa, M., Schell, D., and Twigg, D. (2013a). "Pattern recognition as an indicator of diagnostic expertise," in *Pattern Recognition - Applications and Methods*, Vol. 204, eds P. Latorre Carmona, J. S. Sanchez, and A. L. N. Fred (New York, NY: Springer-Verlag), 1–11.

Loveday, T., Wiggins, M. W., Harris, J. M., O'Hare, D., and Smith, N. (2013b). An objective approach to identifying diagnostic expertise among power system controllers. *Hum. Fact.* 55, 90–107. doi: 10.1177/0018720812450911

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., and Schell, D. (2013c). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Hum. Fact.* 55, 125–137. doi: 10.1177/0018720812448475

Loveday, T., Wiggins, M. W., and Searle, B. J. (2014). Cue utilization and broad indicators of workplace expertise. *J. Cogn. Eng. Decis. Mak.* 8, 98–113. doi: 10.1177/1555343413497019

Matthews, G., and Davies, D. R. (1998). "Arousal and vigilance: the role of task factors," in *Viewing Psychology as a Whole: The Integrative Science of William N. Dember*, eds R. B. Hoffman, M. F. Sherrick, and J. S. Warm (Washington, DC: American Psychological Association).

Moore, C. G., and Müller, S. (2014). Transfer of expert visual anticipation to a similar domain. *Q. J. Exp. Psychol.* 67, 186–196. doi: 10.1080/17470218.2013.798003

Morrison, B., Wiggins, M. W., Bond, N., and Tyler, M. (2013). Measuring cue strength as a means of identifying an inventory of expert offender profiling cues. *J. Cogn. Eng. Decis. Mak.* 7, 211–226. doi: 10.1177/1555343412459192

Moutafi, J., Furnham, A., and Tsaousis, I. (2006). Is the relationship between intelligence and trait neuroticism mediated by test anxiety? *Pers. Individ. Differ.* 40, 587–597. doi: 10.1016/j.paid.2005.08.004

Müller, S., and Abernethy, B. (2012). Expert anticipatory skill in striking sports: a review and a model. *Res. Q. Exerc. Sport* 83, 175–187. doi: 10.1080/02701367.2012.10599848

Müller, S., Abernethy, B., and Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler's intention? *Q. J. Exp. Psychol.* 59, 2162–2186. doi: 10.1080/02643290600576595

Neerincx, M. A., and de Greef, H. P. (1998). Cognitive support: extending human knowledge and processing capacities. *Hum. Comput. Interact.* 13, 73–106. doi: 10.1207/s15327051hci1301_3

Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., and Parasuraman, R. (2014). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage* 85, 909–917. doi: 10.1016/j.neuroimage.2012.11.061

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation: Advances in Research*,

Vol. 4, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York, NY: Plenum), 1–17.

Oltman, P., Raskin, E., and Witkin, H. (2003). *Group Embedded Figures Test.* California, CA: Mind Garden, Inc.

Pauley, K., O'Hare, D., and Wiggins, M. (2009). Measuring expertise in weather-related aeronautical risk perception: the validity of the Cochran–Weiss–Shanteau (CWS) Index. *Int. J. Aviat. Psychol.* 19, 201–216. doi: 10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., and Fogarty, G. (2013). The application of reduced-processing decision support systems to facilitate the acquisition of decision-making skills. *Hum. Fact.* 55, 535–544. doi: 10.1177/0018720812467367

Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section* 1. General Overview. San Antonio, TX: Harcourt Assessment.

Raven, J., Raven, J. C., and Court, J. H. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 3: The Standard Progressive Matrices.* San Antonio, TX: Harcourt Assessment.

Rose, C. L., Murphy, L. B., Byard, L., and Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. *Eur. J. Pers.* 16, 185–200. doi: 10.1002/per.451

Salthouse, T. A. (1991). "Expertise as the circumvention of human information processing," in *Toward a General Theory of Expertise: Prospects and Limits*, eds K. A. Ericsson and J. Smith (Cambridge, NY: Cambridge University Press), 286–300.

Salvendy, G., and Smith, M. J. (eds) (1981). *Machine Pacing and Occupational Stress.* London: Taylor & Francis.

Scerbo, M. W., Greenwald, C. Q., and Sawin, D. A. (1993). The effects of subject-controlled pacing and task type on sustained attention and subjective workload. *J. Gen. Psychol.* 120, 293–307. doi: 10.1080/00221309.1993.9711149

Schriver, A. T., Morrow, D. G., Wickens, C. D., and Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Hum. Fact.* 50, 864–878. doi: 10.1518/001872008X374974

Schvaneveldt, R. W., Beringer, D. B., and Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *Int. J. Aviat. Psychol.* 11, 253–280. doi: 10.1207/S15327108IJAP1103_02

Schyns, P. G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. *Cognition* 67, 147–179. doi: 10.1016/S0010-0277(98)00016-X

Simonton, D. K. (2007). Talent "and" expertise: the empirical evidence for genetic endowment. *High Ability Stud.* 18, 83–84. doi: 10.1080/13598130701350890

Simonton, D. K. (2008). Scientific talent, training, and performance: intellect, personality, and genetic endowment. *Rev. Gen. Psychol.* 12, 28–46. doi: 10.1037/1089-2680.12.1.28

Singer, R. N., and Janelle, C. M. (1999). Determining sport expertise: from genes to supremes. *Int. J. Sport Psychol.* 30, 117–150.

Small, A. J., Wiggins, M. W., and Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Appl. Cogn. Psychol.* 28, 481–487. doi: 10.1002/acp.3016

Smeeton, N. J., Ward, P., and Williams, A. M. (2004). Do pattern recognition skills transfer across sports? A preliminary analysis. *J. Sports Sci.* 22, 205–213. doi: 10.1080/02640410310001641494

Smit, A. S., Eling, P. A., and Coenen, A. M. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychol.* 115, 35–42. doi: 10.1016/j.actpsy.2003.11.001

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., and Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Hum. Factors* 42, 183–194. doi: 10.1518/001872000779656480

Thompson, C. P., Cowan, T. M., and Frieman, J. (1993). *Memory Search by a Memorist.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Vessey, I., and Galletta, D. (1991). Cognitive fit: an empirical study of information acquisition. *Informat. Syst. Res.* 2, 63–84. doi: 10.1287/isre.2.1.63

Vickers, J. N. (1996). Visual control while aiming at a far target. *J. Exp. Psychol.* 22, 342–354.

Vickers, J. N. (2007). *Perception, Cognition, and Decision Training: The Quiet Eye in Action.* Champaign, IL: Human Kinetics.

Vickers, J. N., and Williams, A. M. (2007). Performing under pressure: the effects of physiological arousal, cognitive anxiety, and gaze control in biathlon. *J. Motor Behav.* 39, 381–394. doi: 10.3200/JMBR.39.5.381-394

Weiss, D. J., and Shanteau, J. (2003). Empirical assessment of expertise. *Hum. Fact.* 45, 104–116. doi: 10.1518/hfes.45.1.104.27233

Wiggins, M., and O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: a cross-sectional analysis of general aviation pilots. *J. Exp. Psychol.* 1, 305–320.

Wiggins, M., Stevens, C., Howard, A., Henley, I., and O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision making. *Austr. J. Psychol.* 54, 162–167. doi: 10.1080/00049530412331312744

Wiggins, M. W. (2014). The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theor. Issues Ergon. Sci.* 15, 283–292. doi: 10.1080/1463922X.2012.724725

Wiggins, M. W. (2015). *Diagnostic Expertise in Organizational Environments.* Surrey: Ashgate Publishing.

Wiggins, M. W., Brouwers, S., Davies, J., and Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise. *Front. Psychol.* 5:541. doi: 10.3389/fpsyg.2014.00541

Wiggins, M. W., Harris, J., Loveday, T., and O'Hare, D. (2010). *EXPERT Intensive Skills Evaluation (EXPERTise) Test.* Sydney: Macquarie University.

Williams, A. M., Singer, R. N., and Frehlich, S. G. (2002). Quiet eye duration, expertise, and task complexity in near and far aiming tasks. *J. Mot. Behav.* 34, 197–207. doi: 10.1080/00222890209601941

Witkin, H. A. (1976). "Cognitive style in academic performance and in teacher-student relations," in *Individuality in Learning*, eds S. Messick, et al. (San Francisco, CA: Jossey-Bass), 38–72.

Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. (1971). *A Manual for the Embedded Figures Tests.* Palo Alto, CA: Consulting Psychologists Press.

Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. (2002). *Group Embedded Figures Test. Sampler Set. Manual and Sample Figures and Scoring.* Menlo Park, CA: Mind Garden Inc.

Wulf, G., McNevin, N. H., Fuchs, T., Ritter, F., and Toole, T. (2000). Attentional focus in complex skill learning. *Res. Q. Exerc. Sport* 71, 229–239. doi: 10.1080/02701367.2000.10608903

Xiao, Y. M., Wang, Z. M., Wang, M. Z., and Lan, Y. J. (2005). The appraisal of reliability and validity of subjective workload assessment technique and NASA-Task Load Index. *Chin. J. Industr. Hygiene Occupat. Dis.* 23, 178–181.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory Pract.* 41, 64–70. doi: 10.1207/s15430421tip4102_2

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *Am. Educ. Res. J.* 45, 166–183. doi: 10.3102/0002831207312909

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Macrocognition in Day-To-Day Police Incident Response

*Chris Baber\* and Richard McMaster*

*School of Engineering, University of Birmingham, Birmingham, UK*

Using examples of incidents that UK Police Forces deal with on a day-to-day basis, we explore the macrocognition of incident response. Central to our analysis is the idea that information relating to an incident is translated from negotiated to structured and actionable meaning, in terms of the Community of Practice of the personnel involved in incident response. Through participant observation of, and interviews with, police personnel, we explore the manner in which these different types of meaning shift over the course of incident. In this way, macrocognition relates to gathering, framing, and sharing information through the collaborative sensemaking practices of those involved. This involves two cycles of macrocognition, which we see as 'informal' (driven by information gathering as the Community of Practice negotiates and actions meaning) and 'formal' (driven by the need to assign resources to the response and the need to record incident details). The examples illustrate that these cycles are often intertwined, as are the different forms of meaning, in situation-specific ways that provide adaptive response to the demands of the incident.

Keywords: macrocognition, sensemaking, police, incident response

## INTRODUCTION

We consider Police incident response as a form of macrocognition (Klein et al., 2003). The primary research question relates to the manner in which a collection of individuals, a 'community of practice' (Wenger et al., 2002), develop a common understanding of the problem that they are addressing through processes of sensemaking. We propose that sensemaking can be a collaborative activity within a given community of practice. This activity is shaped by the institutional frames of the community of practice, which define the formal and informal rules by which information is defined and shared. These rules can be seen in the manner in which the community of practice manages 'meaning', in its collaborative sensemaking. We consider three types of meaning, which we term 'negotiated' (in which informal, unstructured accounts of the incident are shared and clarified), 'structured' (in which formal accounts are logged), and 'actionable' (in which the commentary on the incident informs decisions on how to resource and manage the response). For this paper, a key issue in macrocognition, therefore, relates to this question of how these different meanings are recognized and managed.

In terms of 'community of practice', the incidents that we consider involve Standard Operating Procedures. This means that there is an established organization of individuals, operating within a well-defined domain, and who "...*share a common set of patterns of interpretation, implicit assumptions, and beliefs*..." (Burnett et al., 2004, p. 12). The manner in which a Community of Practice shares its knowledge and understanding involves what we have previously called Collaborative Sensemaking (Duffy and Baber, 2013), which combines 'semantic' sensemaking (in

which a group of people seek to develop a common interpretation of an event, i.e., determining what is known) and 'pragmatic' sensemaking (in which a group of people can be allocated different roles in terms of holding or sharing information, i.e., determining who knows what).

From the point of view of 'pragmatic sensemaking', a Community of Practice shares information partly through common jargon (and associated experience and 'world view') and partly through shared communication technologies and practices. An irony of this (for the type of incident response considered in this paper) is that 'outsiders' (i.e., people who are not part of the Community of Practice) are the very focus of its activity. One implication of this is that there is a need to develop and manage a wide range of 'interfaces' between the Community of Practice and those outside it. These interfaces could be formal, e.g., in terms of Press conferences or briefings to politicians, or informal, e.g., in terms of reassuring members of the public. Central to these interfaces is the need to define the 'meaning' of an incident at the most appropriate level of detail.
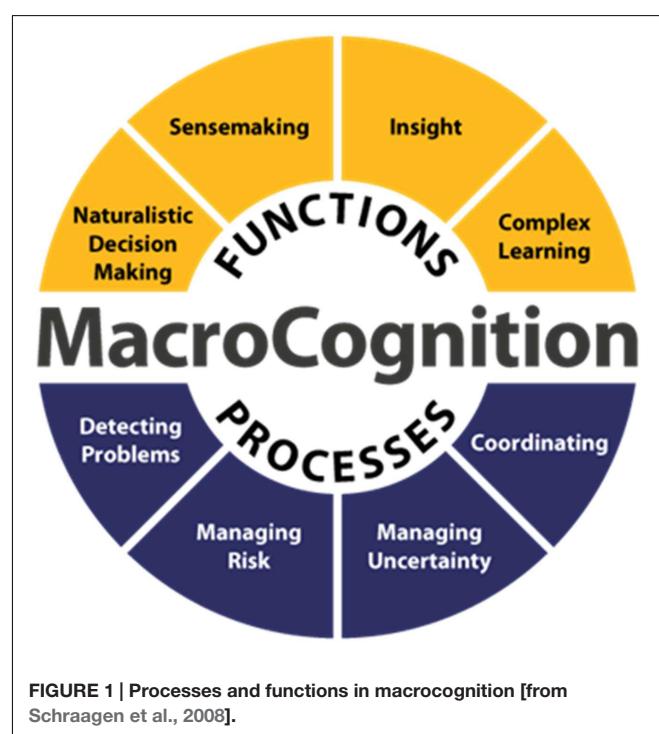
The information sources provide frames (Klein et al., 2006a,b, 2007) for interpreting and responding to the incident. Of particular interest are the institutional frames that are designed to aid the management and recording of incidents, such as the electronic forms that allow call handlers to enter information into incident logs. These electronic forms are a repository of prior experience of the organization; they reflect the primary types of incident to which responses are required and the primary types of information that need to be recorded in order to produce consistent, structured accounts of the incident and the response. In addition to these electronic forms, other types of institutional frame are the policies that local police forces might enact, either in response to National policy or in response to local crime patterns. These policies could emphasize the importance of prioritizing response to some types of incident. Finally, institutional frames could come from the collective experience of the personnel involved in incident response, i.e., the community of practice of incident responders, in terms of expectations of how an incident might develop.

The notion that institutional frames can influence decision making echoes the question posed by Manning (1988), viz. "How does organized rationality interface with the variegated dilemmas and perplexities of human communication?" (p. *xv*). Our reading of this question is in terms of the potential conflict between the Naturalistic Decision Making that personnel involved in incident response will apply and the 'rules' that are embedded in the forms and procedures that they apply. For Manning (1988), these 'rules' might be informal, reflecting concerns of Police Officers, Incident Controllers and Call Handlers (in terms of acceptable ways of behaving on and off duty) and which we see as constituting the community of practice of incident response. Additionally, the 'rules' might be formal and dictate how information is recorded, shared and acted upon, i.e., as institutional frames. From the perspective of macrocognition, this patchwork of 'rules' will influence the space in which information is interpreted, and the ways in which different 'framing' of the same information can vary.

## Incident Response and Macrocognition

Incident response has been extensively researched for major and catastrophic incidents (Dynes, 1970; Quarantelli, 1999; Boin, 2004; Mendonça et al., 2007; Becerra-Fernandez et al., 2008; von Lubitz et al., 2008; McMaster and Baber, 2012). There has been less work on the routine incidents that emergency services face on a day-to-day basis (e.g., Blandford and Wong, 2004, explored Situation Awareness of operators in medical dispatch). Incident response tends to follow a standard process in which a call is received by a Call Handler, responding units are dispatched by the Incident Controller and these units attend and resolve the incident, and the incident in closed. Over the course of this process, an Incident Log is maintained to record relevant information and personnel communicate with each other (via radio) and with members of the public (via telephone or face-to-face).

**Figure 1** illustrates core processes and functions related to macrocognition. The processes {detecting problems, managing risk, managing uncertainty, coordinating} are central to incident response. Indeed, these are the primary processes involved in this activity (the only addition here is the process of managing the Incident Log – which we will argue is an essential part of incident response, not only in terms of recording what has been done but also as part of the coordinating process). In terms of the functions, we will present examples of incident response to show how the situations and prior experiences of personnel involved in the response can exhibit characteristics of Naturalistic Decision Making and Sensemaking. We have less to say on Insight and Complex Learning in this paper (although both can play important roles in the response to incidents and handling of crime).



FIGURE 1 | Processes and functions in macrocognition [from Schraagen et al., 2008].

Central to the activity of the Incident Controller is the need to ensure an optimal resource has been dispatched to the incident: too few officers and there might be a risk to the officers or the public, or they might be unable to apprehend the suspect; too many and there could be problems in resourcing subsequent calls. As Blandford and Wong (2004) note, the decisions governing how to resource a response is as much a matter of situation awareness as it is of policy, and the situation awareness includes not only the location and availability of units which could respond but also the type of response which is required.

## METHODOLOGY

Over the course of 5 years, the second author worked as a Special Constable (volunteer officer) for a Police Force in the UK. During this time he received training on incident response and attended incidents, working 70 shifts in a two-officer patrol crew deployed in a marked police vehicle. These participant observation sessions enabled direct access to the 'on the ground' incident response process, something which is not normally possible for researchers. Informal interviews were conducted with crewmates after incidents had been resolved; notes were taken during patrols whenever possible. These were later supplemented with electronic incident logs for timings and other details. In addition, permission was granted to collect data from the communications centers of two Police forces; over the course of some 30 data collection sessions, we were able to interview and observe Call Handlers and Controllers at work, listen in to 999 (emergency) calls and Police radio traffic, review electronic incident logs. Interviews were done on an opportunistic basis – with questions tailored to clarify the activities that had just been observed. During these interview and observation sessions, data capture was limited to note taking, which ranged from detailed descriptions of activities being undertaken to verbatim recording of telephone and radio conversations.

Such access resulted in a wealth of material. However, this leads to the inevitable problem of deciding what material to select and report. While it is tempting to select those incidents in which there is some level of excitement or novelty, this does not reflect day-to-day operations. On the other hand, some of the more common incidents reveal little of interest about the nature of incident response. For example, a spate of incidents in which gardening equipment was stolen from sheds in back yards might take up a sizeable portion of time but does not make for interesting reading. Typical examples of day-to-day incidents include:

- Burglaries in progress;
- Criminal damage (including arson);
- Domestic violence;
- Medical emergencies (including suicidal and acute mental health problems);
- Retail thefts;

- Road traffic incidents
- Serious assaults;
- Street robberies;
- Urgent welfare concerns (e.g., elderly and disabled persons collapsed in their homes);
- Vehicle crime (e.g., theft from, theft of and driving offenses).

These different types of incident present a range of challenges and risks to the public and responding Officers. Thus, the type of incident will dictate the approaches that are used to respond to them (Flin et al., 2007). For this paper, we have selected a set of incidents that reflect the need for immediate attendance with the opportunity of arresting the suspect (burglaries in progress) or the need to attend the scene to provide assistance (street robbery). We make no claims as to how representative these incidents are of day-to-day policing; we estimate that such incidents would occur three or four times a week, rather than daily, but they represent examples of incidents that those involved would recognize as common. Furthermore, we have chosen not to report incidents which involve violence to the person or domestic violence that contains details which are harrowing and difficult to read.

We present the incidents in two ways. The first is through the use of short vignettes, in which excerpts from incident logs, or verbatim transcripts of radio traffic, are taken from a single incident. The incident transcripts reflect as much information as we feel is necessary for the reader to appreciate what is being discussed or recorded, while also respecting the need to maintain a degree of anonymity in the recorded information. The second is in the form of graphical depictions, which represent the distillation and interpretation of multiple observations and thus are general descriptions of the macrocognitive activities being described. These presentations complement one another, with the vignettes helping the reader to view the diagrams, which in turn provide a framework within which the activity described in the vignettes takes place.

Given the opportunity to collect data in this manner, it is appropriate to ask whether alternative approaches could have been feasible or produced more reliable data. We opted for a participant observation and interview-based approach, with the primary focus on the Police officers and associated staff and the processes that they follow. This means that, in comparison with an ethnographic approach (Hammersley and Atkinson, 1995), this study is heavily prescribed by the information flow and operating procedures. Our descriptions show how information is received, processed and passed around the system. What we are not capturing in detail are the assumptions, attitudes and expectations of the personnel involved (or, for that matter, the members of the public who are the subject of these processes). Thus, while the examples used in the paper involve researchers participating in the social practices under investigation, the process-oriented analysis could miss the rationalization through which the participants continually revise their understanding of the situations they encounter. In other words, we are taking the behavior of participants as indicative of the processes that they follow and then inferring the 'meaning'

that these processes involve. Where practicable we have sought to corroborate our interpretation of meaning with the participants, but the study is not focused on extracting notions of sense and meaning directly from the participants. We believe that the approach taken provides opportunity to triangulate data (through multiple sources of information being collected for each example), investigator (through continued exploration of assumptions made by the two authors in their analysis) and theory (through developing an explanation of the processes that we are observing).

## Coding, Synthesizing, and Representing the Data

For the textual descriptions of the incident, the presentation format is to use CAPITALS to indicate material typed in to the Incident Log (with time of entry on the left), i.e.,

> **14:20** Controller 1: *"THE IP HAS BEEN STRUCK AND FELL TO THE FLOOR"*

and for verbal communications to be presented in *italics*, i.e.,

> Whiskey 3–5: *"Yes – confirmed break-in."*

In the examples in this paper, the textual descriptions are verbatim accounts recorded *in vivo*.

The graphical description was originally developed in McMaster and Baber (2005) and is intended to show how cognitive activity in spread across actors and artifacts. **Table 1** lists features of the activity and which can be used as the basis of a simple task analysis.

The features from **Table 1** are combined into a diagram which shows the flow of information in an incident response (**Figure 2**). The diagram shows the key transformations of information (e.g., from one modality or storage medium to another). Thus, **Figure 2** shows the process through which an Incident Controller (in the first panel) responds to an open Incident Management System (IMS) log for an incident requiring immediate attendance, and then puts out a call to all units to ask for attendance. Of the units that respond, one unit asks for further details on the location. As the incident unfolds, the Incident Controller provides further information relating to access to the property.

While **Figure 2** provides a summary of the incident, we are aware that such a representation is not without its

problems. Any description (verbal or graphical) stands or falls on the comprehensiveness of its content and, consequently, reflects the selectivity of the analyst. As far as practicable, we have included those elements of the incident which were 'external', i.e., available to participants in the incident response, e.g., the content of the Incident Log, verbatim transcriptions of communications over the radio. This means that we have not included the reflections, assumptions, interpretations and other 'internal' elements of the responders. Nor have we provided much in the way of contextual or situational material for each incident. However, we feel that the material that we report is sufficient to allow us to draw conclusions relating to the macrocognition involved in incident response.

The approach to coding of the examples, in terms of type of meaning, is explained for each example. In broad terms, where participants are asking questions or where there is evidence of confusion, we consider this to be negotiated meaning. Here, the participants are, we believe, seeking to establish common ground in order to make sense of the incident. Where participants are giving direct instructions, we consider this to be actionable meaning. Here, the participants are either providing information or instructions that enable other participants to effect an action. Where information is being typed in to the Incident Log, we consider this to be structured meaning. Here, the information is being formatted for subsequent use. As the examples illustrate, this distinction is not always clear-cut; information might be structured (in the sense that it is typed in to the Incident Log) but could also involve negotiation, with participants raising questions or debating the meaning of the information. We also note that several of the examples show overlap in the types of meaning, i.e., it is rarely the case that the incident proceeds with negotiated meaning at the start, leading to actionable meaning and then to structured meaning in the final report. Rather, the incidents appear to shift between these meaning types.

## THE 999 CALL: FROM NEGOTIATED TO ACTIONABLE MEANING

Whalen and Zimmerman (1990) show Callers often present imprecise and hesitant openings to their calls. As Baber et al. (2006) point out, rather than taking a verbatim account of the Caller's information, the Incident Controller will translate this information into a format which is more suited to the structure of the Incident Log. Thus, one of the roles of the Call Handler is to negotiate the meaning of the incident with the Caller. This negotiation is supported by a set of core questions that Call Handlers are trained to use in order to direct the conversation and to establish the important facts quickly, e.g.,

- Call Handler: *"Have you been injured?"*
- Call Handler: *"Where did they go?"*

Caller, Call Handler, and Incident Controller develop some form of common ground (Clarke, 1996). In this concept,

**TABLE 1 | Features of Activity.**

| Feature | Description |
| --- | --- |
| Agents | Who is involved (people/artifacts) |
| Activity | The purpose of the operation |
| Modality | Information state (verbal, text, etc.) |
| Form | Language style, abbreviations, etc. |
| Transmission | How is information shared |
| Transformation | How is information acted upon |
| Storage | How is information retained |
| Resource for action | Actions cued by representation |

common ground is *"the mutual knowledge, beliefs, and assumptions shared by the speaker and addressees."* (Clarke, 1996, p. 247). Clarke's (1996) concept of common ground proposes that people draw on three sources of information:

- Perceptual evidence (the experience to which people have access);
- Linguistic evidence (the words that people are hearing);
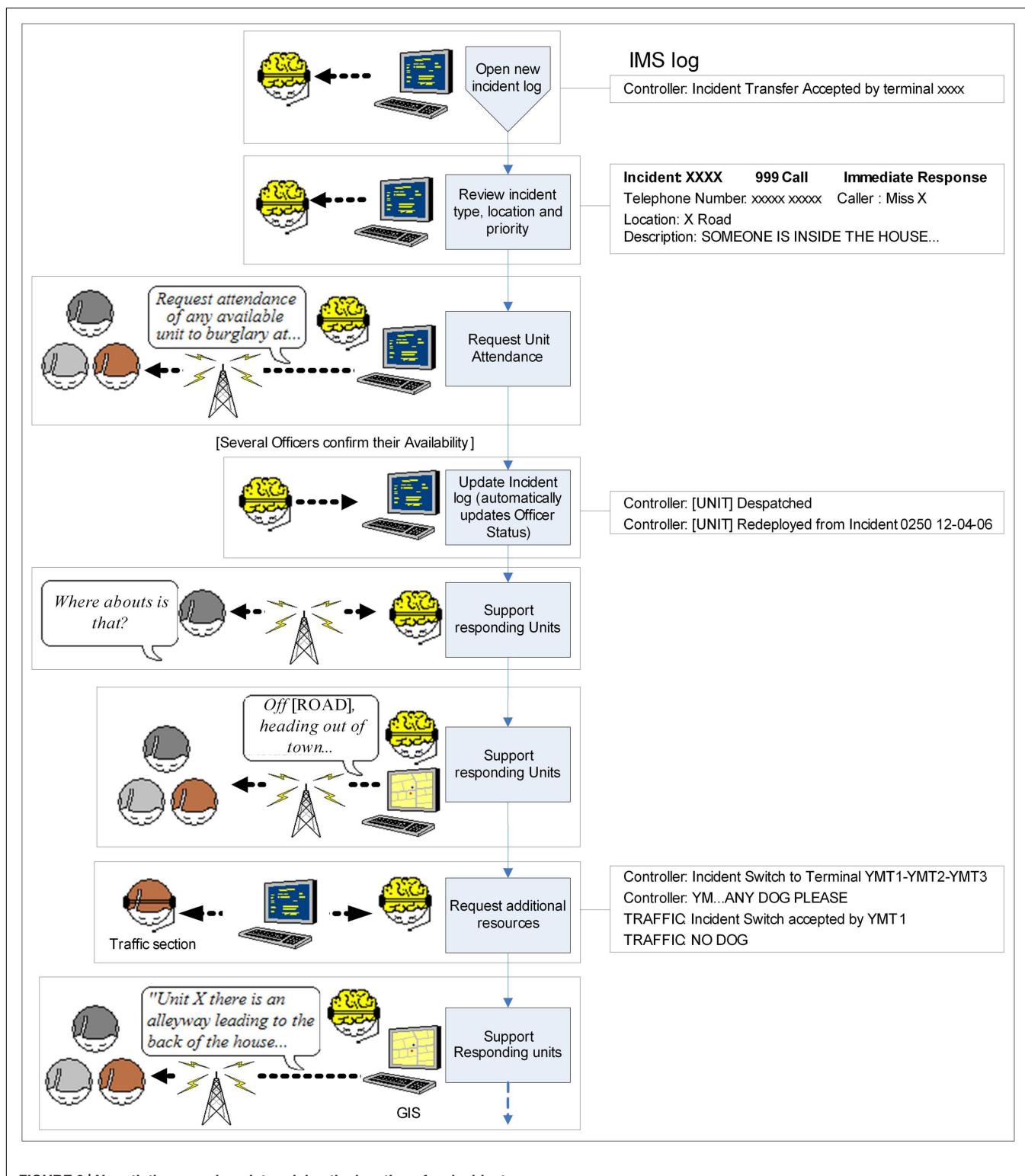


**FIGURE 2 | Negotiating meaning: determining the location of an incident.**

● Community evidence (knowledge which they might believe is shared within a given community, perhaps as the result of training or enculturation).

The Caller, Call Handler and Incident, Controller will not have the same perceptual evidence (the Call Handler and Incident Controller are removed from the scene that the Caller is witnessing or recalling). Thus, part of the conversation is aimed at translating the Caller's perceptual evidence into actionable meaning (to support the Incident Controller) and part of the conversation is aimed at translating the Caller's perceptual evidence into structured meaning (to support completion of the Incident Log). In terms of linguistic evidence, a key role of the Incident Controller is to translate the words of the Caller into the terminology used by the Police. For instance, the description of an offender may change from *"white lad"* to *"IC1 male"*, which is the relevant UK Police National Computer Ethnicity Classification. Abbreviations and acronyms are also employed, for example *"My car has been stolen"* is formalized within the Police as *"Theft of Motor Vehicle"*, which is written as *"TOMV"*. This terminology and jargon relates to community evidence. Furthermore, one might find Caller's seeking to provide information in a manner which they believe fits the community knowledge of the Police, e.g., when reporting a car's registration number, the Caller might use the ICAO (International Civil Aviation Organization) alphabet of A, alpha; B, bravo; C, Charlie etc. because they believe that this is how to report letters of the alphabet to a Police Officer. Of course, the Caller's might not know all of the words used in the ICAO alphabet and so might use their idiosyncratic versions, such as A, apple; B, baby etc., but the intent of providing clear definition of letters over a potentially noisy communication channel remains the same.

For us, common ground represents the meaning that the Caller and Call Handler are negotiating during the call, and which is then translated into structured meaning by the Call Handler to record onto the IMS so that it could be read by an Incident Controller, interpreted in terms of actionable meaning, and subsequently communicated over the radio to responding units. In terms of macrocognition, 'common ground' implies the need for a community of practice to work within its institutional frames to gather the appropriate 'community evidence', i.e., that information is selected which corresponds to working practices and which has been recorded in an acceptable. The processes by which information is selected and recorded relate to our notions of 'meaning'. **Figure 2** illustrates some of the issues surrounding common ground in incident management. In response to an initial call, the Incident Controller issues a 'Request for attendance' to the incident at 'x road'. The first response to this request is to ask 'where abouts is that?' to which the Incident Controller provides further geographical information. Here, the relevant information is being explored and elaborated in a form of negotiated meaning. As Attending Officers reach the address, the Incident Controller provides further information about the geography ('. . .an alleyway leading to the back of the house. . .').

In this example, the unfolding activity can be seen as the effort after actionable meaning, i.e., to provide sufficient information to the Attending Officers to allow them to operate at that address.

The incident summarized in **Figure 2** has an Incident Log entry of "no dog", indicating that it is not possible to supply a police dog to this call. In order to make these decisions, the Incident Controller draws on the incident classification made by the Call Handler in response to the original call and recorded the Incident Log. Often the classification (and required response) is negotiated through the editing of the Incident Log as the response unfolds. What we find particularly interesting is the decision of what to include in the Incident Log; when the Incident Controller (and Call Handler) speaks to Attending Officers, members of other services or members of the public, what is recorded is not a verbatim account but as accurate a gist of the conversation as is sufficient for the log. At this level, macrocognition applies to the translation of negotiated meaning (i.e., the content of conversations which might require clarification) into structured meaning (i.e., which can be written into the Incident Log).

**Figure 3** summarizes the process of taking a 999 call arising from a street robbery. The boxes on the right-hand side of the figure show the information that is recorded in the notepad and incident log at various points, showing how the incident log gradually develops during the course of the call. The figure also illustrates how the log structures the incident details and mediates indirect communications between the Call Handler and Control (the Call Handler can see that Control has dispatched a unit to the incident and is able to tell the Caller that the Police will be with them soon).

In the incident summarized in **Figure 3**, the caller provides initial information about the incident, i.e., "My boyfriend has been mugged...Two lads...they took his mobile phone." While this provides some information about the nature of the incident, it does not provide information that might be relevant to the response, such as whether any injuries had been sustained. Thus, the initial call log records a location and a likely destination for the perpetrators, i.e., 'x school'. Again, the aim is to provide sufficient actionable meaning for the response to be made.

In the following extract, two Incident Controllers are jointly handling multiple incidents on the same radio talk group; they update the same Incident Log relating to the ongoing reporting of a violent robbery. The timestamps (minutes and seconds since the start of the call, on the left of the text) indicate when information is typed into the log; where there are gaps in the timestamps, e.g., 14:27 to 14:58, this is likely to be where one of the Controllers is speaking with the Officer Attending. In this log, two issues are raised and resolved. The first issue involves concerns with the victim Injured Party (IP) of an attack and whether or not an ambulance (Ambo) is needed. The second concerns the need for Scene of Crime Officers (SOCO) to attend the scene to gather evidence (which involves notifying a third controller).

**14:20** Controller 1: *"THE IP HAS BEEN STRUCK AND FELL TO THE FLOOR"*
**14:27** Controller 1: *"OFFICERS CHECKING TO ASCERTAIN IF AMBO REQUIRED."*
**14:58** Controller 2: *"LADY HAS BEEN KNOCKED OVER AT DOOR WHEN OFFENDERS"*
**15:00** Controller 2: *"GAINED ENTRY"*
**15:15** Controller 1: *"CAN SOCO ATTEND ASAP PLSE"*
**15:30** Controller 1: *"FROM OFFICERS THE FEMALE IP DOES NOT REQUIRE AMBO AS"*
**15:32** Controller 2: *"PLS GET SOCO FOR THIS"*
**15:39** Controller 1: *"IP STATES HAS NO INJURIES"*
**16:06** Controller 2: *"ASKING FOR AMBO ELDERLY FEEMAL BADLY SHAKEN APPROX"*
**16:07** Controller 2: *"86 YRS"*
**16:23** Controller 3: Incident Accepted
**17:36** Controller 1: *"THE OFFICERS NOW ASKING FOR AMBO AS THE IP 86YRS OLD"*
**17:51** Controller 1: *"IS EXTREMELY DISTRESSED-UPSET"*
**17:54** Controller 2: *"AMBO LOG [Number]"*

**18:34** Controller 3: *"SOCO INFORMED."*
**18:40** Controller 3: This incident added to SOCO list for section [Number]

It is noteworthy that in this example there is no spoken communication between the three Incident Controllers, two of whom are co-located. Rather, the updating the Incident Log provides the development of common ground concerning the incident. Thus, Controller 1 (15:30) suggests that an ambulance is not required but subsequently Controller 2 (16:06) disagrees and requests an ambulance. Both entries are made in response to comments from the Attending Officers (as indicated in the Incident Log) and, so the change in entries reflects changes in the assessment of the situation made at the scene. For Controller 1, there was no need for the ambulance as the Attending Officers report that the IP "has no injuries" (15:39) but for Controller 2, they note the age of the IP and the she is "badly shaken" (16:06). Controller 1 then also logs the request for ambulance as the IP is "extremely distressed – upset" (17:51). Controller 2 logs the request for an ambulance to attend (17:54). In this example,
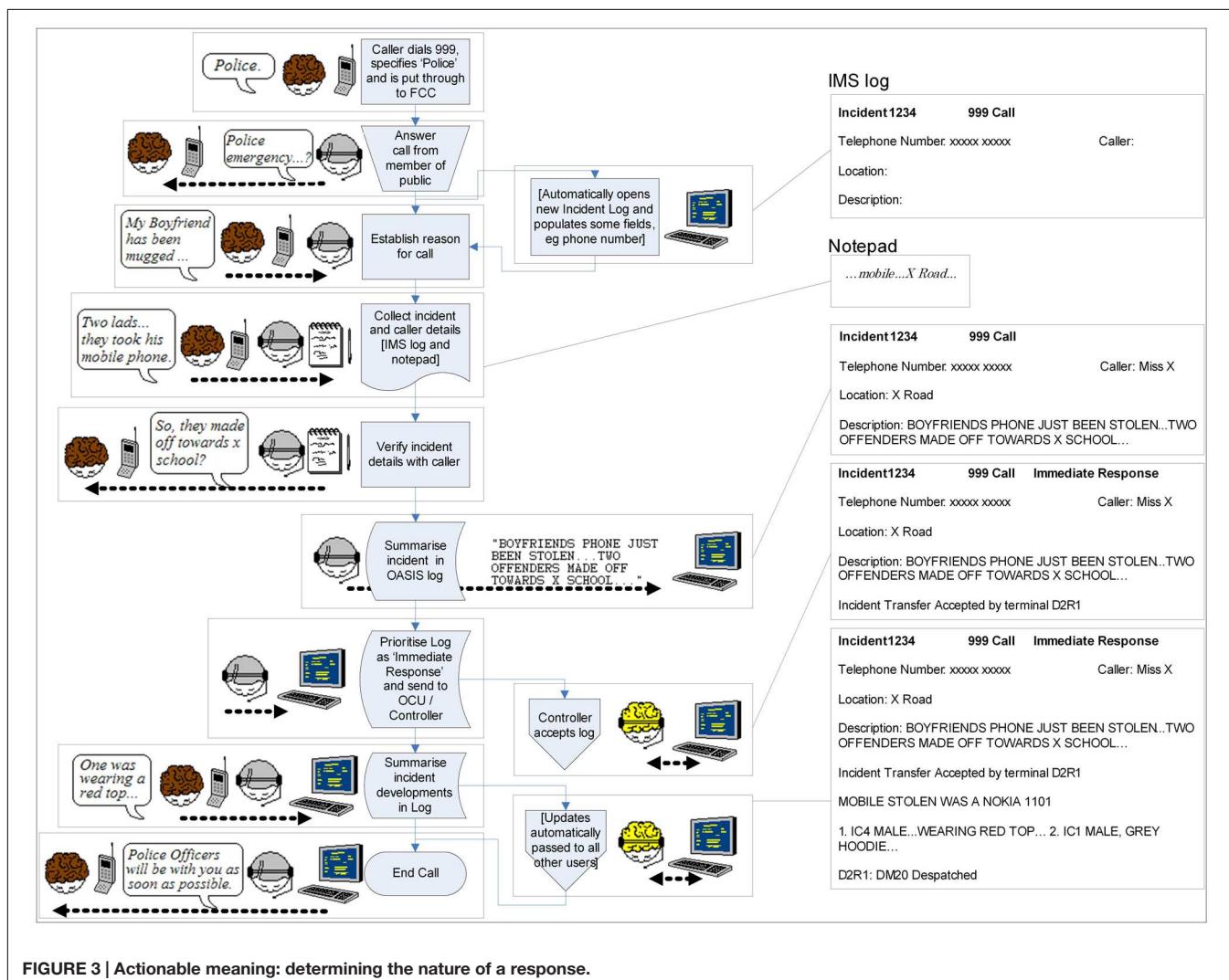


**FIGURE 3 | Actionable meaning: determining the nature of a response.**

the updating of the Incident Log provides both a record of the management of the incident, i.e., structured meaning, that could provide the basis for subsequent enquiries (i.e., the condition of the victim could be used as part of the prosecution against the perpetrator) and negotiated meaning, i.e., in terms of deciding whether or not to call for an ambulance. The example concludes with actionable meaning, i.e., an ambulance is called and a Scene of Crimes Officers is tasked with visiting the scene. What is particularly pertinent about this example is that way in which the three types of meaning are interspersed, and the way in which the negotiation is performed through comments on the Incident Log rather than through verbal communication (even, as we have already noted, two of the Incident Controllers are adjacent to each other in the control room).

## ATTENDING OFFICERS TRAVELING TO THE INCIDENT

As they make their way to the incident, Officers plan their response in terms of risk (threat assessment), powers and policy, and tactics. For Borglund and Nuldén (2008) this represents a form of 'active traveling', in which the Officers will not only search the streets at they drive for vehicles or persons of interest and for the address to which they have been directed, but also review experience of previous, similar calls. Although the Officers will have received some initial details from the Controller, these are often only the bare minimum, such as an approximate location and a statement of the nature of the incident, for example *"male being assaulted by two males"*. In terms of the macrocognitive process of 'managing risk' (**Figure 1**), the first indication of the level of risk associated with the incident (both to members of the public and the responding Officers) and consequently the appropriate response, will come from the type of incident. When an offender is named by the Caller, Officers might ask the Controller to run a check through the Police National Computer; if the person is known to the police, this will provide a summary of any previous arrests or convictions, as well as warning markers (i.e., drugs, violence, weapons or self-harm) associated with those individuals.

If the Call Handler updates the log as a result of further conversation with the Caller (e.g., description of an offender, their direction of travel, vehicle, etc.), this information will be visible to the Controller, who passes this to the Officers. As a result of these further updates, the responding units may change their tactics, for example, if the offender has left the scene Officers may decide to perform a search of the area before speaking to the victim, in the hope of catching the offender.

In terms of the macrocognitive processes of 'managing uncertainty' and 'detecting problems' (**Figure 1**), Attending Officers may ask the Controller to check IMS for: previous emergency calls to that location, details of any persons associated with that location and any previous convictions or warning markers (e.g., for violence or weapons) associated with those individuals. For example, the IMS will indicate if previous 999 calls have been made from a number, or if any persons named in a log are associated with previous incidents at that address. In

their analysis of Mobile Data Terminal (MDT) use, Branaghan et al. (2010) identified five main clusters of information which could inform decision making of Attending Officers: Potentially Violent, Citizen Welfare, Medical, Traffic, Non-violent. These relate to the macrocognitive demands related to managing uncertainty, managing risk and detecting problems, and could the subject of discussion between Attending Officers and Control, or amongst Officers in a talk group.

Officers will often rely on Controllers to remind them of incident details that they have forgotten – such as house numbers, names, or vehicle registration numbers – radioing the Controller as they near the scene to request that that information is repeated. In **Figure 2**, an Officer asks for some clarification of where the incident location was, and then asks for the name of the company to be repeated. The Controller has pro-actively checked the location using the GIS (Geographical Information System) and, unprompted, provides information to clarify the incident location, i.e., in terms of the 'alleyway' to the back of the house.

## ATTENDING OFFICERS AT THE SCENE

As they arrive at the scene, responding Officers notify the Controller (who updates the incident log); the Officers may be confronted by an ongoing incident, or they may find that the immediate threat from the incident has stopped. Their response to the incident is concerned with: (i) controlling and resolving the situation, and (ii) performing an initial investigation of the events surrounding the incident. Where more than one Officer is deployed to an incident, they may decide to separate and divide tasks between them (e.g., conducting searches, separating belligerent parties, speaking to witnesses), using their radios in point to point mode (i.e., direct one to one) to coordinate their activities without taking up airtime on the talk group.

In terms of the macrocognitive process of 'managing uncertainty' (**Figure 1**), responding to incidents is complicated by the fact that many of the incident details may well be inaccurate, including the caller's account of events, the names or descriptions of parties involved and very often the nature of the incident itself (i.e., the frame selected by the Call Handler during the initial call). In the following example, multiple units respond to reports of a break-in in progress at night; Officers are on the scene within 3 min, however, on their arrival, the property and surrounding houses appear to be secure and undisturbed, casting doubt on the nature of the incident. The Controller switches the incident log back to the Call Hander (in a different Control Room) to double check the address. The situation Officers encounter at the scene is at variance to the summary they have been given, which, in turn, cues activity from the Controllers and Call Handler, who communicate with each other via the IMS (12:46 to 13:28).

**12:46** Controller A:  *"CAN YOU CONFIRM x RD OR x ST"*
**13:00** Call Handler:  "STANDBY"
**13:23** Call Handler:  *"I HAVE LISTEND TO TAPE AGAIN IT IS x STREET"*
**13:28** Call Handler:  *"NOT ROAD - MY APOLOGIES"*

**13:28** Controller A: [Receives no reply from caller's mobile phone.]

**14:10** Controller B: [Updates caller details to x Street]

**14:16** Controller A: [Updates incident location to x Street]

**15:20** Controller B: [Notes that the house numbers in x Street only go up to 12 – the caller had reported living at number 15]

**15:50** Controller B: [Performs searches for the caller on the Electoral Role database]

**17:20** Call Handler: "*I HAVE LISTEN TO ALL THE TAPE AND WHEN I CONFIRM*"

**17:34** Call Handler: "*THE NUMBER OF THE ADDRESS CALLER STATES x ROAD*"

**17:44** Call Handler: "*I REPEATED IT TO HIM AND HE SAID YES x ROAD*"

**17:54** Call Handler: "*AT THE BEGINNING OF THE TAPE HE STATES x*"

**18:06** Call Handler: "*STREET*"

**21:50** Controller B: [Notes that Officers have checked the front and rear of both 12 x St and 12 x Rd and spoken to resident at 12 x St – all in order.]

**24:50** Controller B: [All units are leaving the scene. The log is closed, having been redefined as a false call.]

While we have presented the types of meaning as related to common ground, this does not guarantee that all communications are correct or complete. The notion of negotiated meaning that we are developing in this paper suggests that it is possible for a community of practice to carry more than one interpretation of a situation. These multiple interpretations could arise from problems with the structured information, e.g., in the previous example, the problem was whether the address was 'road' or 'street'. As soon as it became apparent that the incident could not be resolved, it was closed as a 'false call'. In this example, the 'common ground' was not necessarily agreement on the address so much as agreement on the nature of the call (and how to respond to it).

In exceptional circumstances, the talk group becomes an open forum for a group of responding Officers to collaboratively make sense of an incident. This extract shows part of the radio communications during the response to a 'break-in in progress' (burglary), where several Officers were already at the scene, searching for the offender and other resources were en route. As can be seen, Officers are using the talk group to directly communicate in order to coordinate their response, with the Controller playing an ancillary, rather than leading role. Interestingly, although the Sergeant involved provides some leadership to the other units – for example directing units during the search – none of the units involved in the example is demonstrably 'in charge' of coordinating the response. Instead, the units involved jointly make sense of and determine the response to the incident (break-in in progress) and the situation as they find it. This also shows that the Controller has to repeat the incident

details several times, either because a new unit has become involved (Dog Handler), or because details have been forgotten (Whiskey 2).

W3–5 [Sergeant]: "[OFFICER A]: *you're on the wrong side. . ..Unit looking at me, go down there.*"

Dog Handler: "*You were calling me?*"

Control: "*Possible Break in progress...*" [Gives details]

Dog Handler: "*Can you confirm I'm required?*"

Whiskey 3–5: "*Yes – confirmed break-in.*"

. . .

Whiskey 2: "*Whiskey 2: What's the address again?*"

Control: "*[ADDRESS]*"

[Confusion ensues over the location of the road and property]

Control: "*On mapping, you have got* [ROAD]..."

Officer A: "*I'm by* [LOCATION], *is that right?*"

Officer B: "*No, it's further round, near the church....do a left there.*"

Officer C: "*[OFFICER C] to 3-5.*"

Whiskey 3–5: "*Go on.*"

Officer C: "*Can you speak to the IP and see if a laptop's been stolen?*"

Whiskey 3–5: "*Confirmed.*"

Officer C: "*I've found a laptop cable...*"

Whiskey 3–5: "*Does that give a direction of travel?*"

Officer C: "*It goes to a dead end...*"

Whiskey 1: "*Whiskey 1 to Control?*"

Control: "*Go ahead.*"

Whiskey 1: "*Another property is open,* [OFFENDER] *may still be inside.*"

Whiskey 3–5: "[Requests location of this address]"

Whiskey 1 "*...outside IP's address, go back...2nd right...*"

In this example, the different threads of conversation show interconnections between different types of meaning. The negotiated meaning develops over the course of response, e.g., in terms of tasking ('can you confirm I'm required?', 'see if a laptop's been stolen?') and in terms of location ('what's the address gain?', 'is that right?', 'it goes to a dead end. . .'). Incident Controller (Control) is providing information to Attending Officers, in the form of the specific location of the incident. The Attending Officers are sharing information with Control ('another property is open [Offender] may still be inside'). This example captures some of the confusion of incident response, with the need to define the required information to support the response, and the manner in which response can develop as new opportunities arise. The multi-threading of meaning in this example shows how the *ad hoc* planning of incident response creates opportunities to develop common ground between the community of practice. It also provides an interesting insight into the challenges of defining what information to record in the Incident Log, i.e., when to convert the information to structured meaning.

# CLOSING THE INCIDENT

Once the incident has been resolved, the Officer will radio the Controller with a final update that summarizes their assessment of the incident and the actions taken. This narrative could be as short as *"One under arrest for drunk and disorderly – transporting to Custody"*, but may be more lengthy for complex incidents. The Controller will add this final update to the incident log, which is then closed.

# DISCUSSION

We began this paper with the proposal that sensemaking, as collaborative activity, is performed within a given community of practice, operating with the constraints of its institutional frames (which are both formal and informal rules of that community and the technology used to support its activity). In the examples presented in the paper, the rules are instantiated through the ways in which the community of practice manages meaning. As the examples show, the management of meaning is not a neat, linear process but involves the participants raising questions, seeking clarification, misinterpreting information and correcting their understanding. We have used the notion of common ground as a lens through which to consider this activity, but it is also illustrates very nicely the cyclical nature of sensemaking in the Data/Frame model (Klein et al., 2006a,b).

In the examples, the manner in which information is communicated influences the ways in which meaning is managed. For the Attending Officers, communication is almost exclusively spoken, either via radio or face-to-face. This means that Attending Officers tend to only know the content of Incident Log when it is read to them by the Incident Controller. In this case, macrocognition applies to the translation from structured meaning to actionable meaning (i.e., from the entries in the Incident Log to advice and instruction for Attending Officers). As Clarke's (1996) notion of common ground implies, the macrocognition of incident response is a continual process of comparing and contrasting across perceptual, linguistic, community evidence. From another perspective, the notion of Distributed Situation Awareness (Stanton et al., 2006) suggests that teams will typically have different views on a situation, with the possibility that their knowledge overlaps in part rather than completely. This suggests that the macrocognition in incident response relates to deciding what knowledge to share and what format to use for its sharing. For example, the response to the "intruder" shining a torch into someone's window involved sharing of knowledge of previous incidents from this address. This 'informal' knowledge could be shared during the briefing prior to patrols leaving the Police Station or could, as in this instance, be shared over the radio. In this instance, the shared knowledge became integral to the response, i.e., "we'll go and have a chat". When this knowledge applies to the response, it is formally recorded in the Incident Log. Otherwise, it remains part of the informal 'rules' that play against the formal rules for recording. The examples also highlight that the interplay of formal and informal is not a simple matter of all entries in the Incident Log having structured meaning, i.e., there are several examples in which the content of the Incident Log is used to challenge other entries; in such cases, the 'formal' (structured meaning) of the Incident Log is replaced with an 'informal' (negotiated meaning). What is interesting here is that such communication can occur even when a more appropriate channel for informal communication is available, i.e., when Incident Controllers are sitting near each other and can simply talk to each other. This suggests that the notions of formal/informal rules, or negotiated/structured meaning are neither rigid concepts to apply to analysis nor necessarily factors consider in the choices that Incident Controllers make.

In terms of limitations of the work, the use of a selection of examples taken from a larger collection could raise accusations of 'cherry-picking' those examples which best support the points that we are making in the paper. It might have been beneficial to report more examples, or to classify a large collection of examples in terms of the issues identified in the paper. We feel that the examples illustrate the individual nature of the incidents that Police will be responding to. This means that collecting more examples might not necessarily allow reduction to specific types, and hence there is a need to consider individual cases. On the other hand, in order to determine whether the unique characteristics of a specific case can be generalized to similar operations, there is a need to extend the set of examples that are explored, and this could be the subject of subsequent work.

In terms of the lessons that these examples, and our analysis of them, might raise, we believe that there are two lines of exploration that could be developed. The first concerns the nature of sensemaking as collaborative activity. Many of the case studies that have been reported since Weick's (1995) pioneering work on sensemaking draw on analyses of discussions and meetings in which groups make sense of the problems that they face. Thus, there seems a strong case to be made for the proposal that collaborative sensemaking follows the elements outlined in this paper. However, the idea that there is an 'informal' sense which can be used to describe and define a situation only covers part of the processes that sensemaking involves. For many situations (and this is often critical in Emergency Response) there is a parallel requirement to produce a 'formal' statement of the response and this requires description of the situation in terms which can be used to justify the use of resources. Baber et al. (2006) describe how narratives are constructed to develop the crime scene investigation from informal sensemaking to formal reporting. The Incident Log, which is a formalized 'in the moment' account of the incident response as a series of time-stamped event updates which reflect the twists and turns of the ongoing sensemaking process that took place during the incident.

The second line of exploration concerns that nature of the technology and work processes followed in Incident Response. As Manning (1988) notes, there is an ongoing tension between the need to record a formal, reliable, and objective account of the response, and the collaborative search after meaning, which seems to arise spontaneously when groups of people engage in sensemaking. One implication of this is the need to manage the 'meaning' of the incident as it unfolds, and to combine this with the management of the incident itself. For us, this implies two

cycles of macrocognition which partially overlap. The first cycle concerns the formal rules which govern the management of the response, e.g., in terms of recording details in the Incident Log and in terms of providing resources for the response. The second concerns the informal rules which govern the operation of the Community of Practice and support the managing of uncertainty and risk as the incident unfolds. It is this overlap between these two cycles of macrocognition which enables adaptability in the ensuing response and which also the need to ensure that the 'formal' rules do not overwhelm the informal rules.

## AUTHOR CONTRIBUTIONS

The paper was written by CB. The examples and diagrams were collected by RM. Analysis and discussion was written by CB.

## REFERENCES

Baber, C., Smith, P., Cross, J., Hunter, J. E., and McMaster, R. (2006). Crime scene investigation as distributed cognition. *Pragmat. Cogn.* 14, 357–385. doi: 10.1075/pc.14.2.14bab

Becerra-Fernandez, I., Xia, W., Gudi, A., and Rocha, J. (2008). "Task characteristics, knowledge sharing and integration, and emergency management performance: research agenda and challenges," in *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM2008)*, Washington, DC, 88–92.

Blandford, A., and Wong, B. L. W. (2004). Situation awareness in emergency medical dispatch. *Int. J. Hum. Comput. Stud.* 61, 421–452. doi: 10.1016/j.ijhcs.2003.12.012

Boin, A. (2004). Lessons from Crisis Research. *Int. Stud. Rev.* 6, 164–194. doi: 10.1111/j.1521-9488.2004.393_2.x

Borglund, E. A. M., and Nuldén, U. (2008). Making sense in proactive police work. *Paper presented at the 31st Information Systems Research Seminar in Scandinavia, IRIS 31*, Åre.

Branagan, R. J., Takamura, J., Palmer, M. T., Hildebrand, E. A., and Sevier, D. C. (2010). Mapping information requirements for police patrol and response to informational displays. *J. Cogn. Eng. Decis. Mak.* 4, 113–128. doi: 10.1518/155534310X12832748852111

Burnett, M., Wooding, P., and Prekop, P. (2004). "Sense Making – underpinning concepts and relation to military decision-making," in *Proceeding of the 9th International Command and Control Research and Technology Symposium Copenhagen*, Washington, DC: CCRP.

Clarke, H. H. (1996). *Using Language.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620539

Duffy, T., and Baber, C. (2013). "Measuring collaborative sensemaking," in *Proceedings of the 10th International Conference on Information Systems for Crisis Response and Management*, Baden-Baden.

Dynes, R. R. (1970). *Organized Behaviour in Disaster.* Lexington, MA: Heath Lexington Books.

Flin, R., Pender, Z., Wujec, L., Grant, V., and Stewart, E. (2007). Police officers' assessment of operational situations. *Policing* 30, 310–323. doi: 10.1108/13639510710753289

Hammersley, M., and Atkinson, P. (1995). *Ethnography: Principles in Practice.* London: Routledge.

Klein, G., Moon, B., and Hoffman, R. R. (2006a). Making sense of sensemaking II: a macrocognitive model. *IEEE Intel. Syst.* 21, 88–92. doi: 10.1109/MIS.2006.100

Klein, G., Moon, B., and Hoffman, R. R. (2006b). Making sense of sensemaking I: alternative perspectives. *IEEE Intel. Syst.* 21, 70–73. doi: 10.1109/MIS.2006.75

Klein, G., Phillips, J. K., Rall, E. L., and Peluso, D. A. (2007). "A data-frame theory of sensemaking," in *Expertise out of Context: Proceedings of the 6th International Conference on Naturalistic Decision Making*, ed. R. Hoffman (Mahwah, NJ: Erlbaum), 113–155.

Klein, G., Ross, K. G., and Moon, B. M. (2003). Macrocognition. *IEEE Intel. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Manning, P. (1988). *Symbolic Communication: Signifying Calls and the Police Response,* Cambridge, MA: MIT Press.

McMaster, R., and Baber, C. (2005). "Integrating Human Factors into Systems Engineering through a distributed cognition notation," in *Proceeding of the IEE People and Systems Symposium: Who are we Designing for?* London. doi: 10.1049/ic:20050454

McMaster, R., and Baber, C. (2012), Multi-agency operations: cooperation during flooding. *Appl. Ergon.* 43, 38–47. doi: 10.1016/j.apergo.2011.03.006

Mendonça, D., Jefferson, T., and Harrald, J. (2007). Collaborative adhocracies and mix-and-match technologies in emergency management. *Commun. ACM*, 50, 44–49. doi: 10.1145/1226736.1226764

Quarantelli, E. L. (1999). *Disaster Related Social Behavior: Summary of 50 Years of Research Findings.* Delaware: University of Delaware.

Schraagen, J. M., Klein, G., and Hoffman, R. R. (2008). "The macrocognition framework of Naturalistic decision making," in *Naturalistic Decision Making and Macrocognition*, eds J. M. Schraagen, L. Militello, T. Omerod and R. Lipshitz (Aldershot: Ashgate), 3–25.

Schraagen, J. M., and van de Ven, J. (2011). Human factors aspects of ICT for crisis management. *Cogn. Technol. Work* 13, 175–187. doi: 10.1007/s10111-011-0175-6

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., et al. (2006), Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics* 49, 1288–1311. doi: 10.1080/00140130600612762

von Lubitz, D. K. J. E., Beakley, J. E., and Patricelli, F. (2008). All hazards approach' to disaster management: the role of information and knowledge management, Boyd's OODA loop, and network-centrality. *Disasters* 32, 561–585. doi: 10.1111/j.1467-7717.2008.01055.x

Weick, K. (1995). *Sensemaking in Organizations.* Thousand Oaks, CA: Sage Publication.

Wenger, E., McDermott, R., and Snyder, W. M. (2002). *Cultivating Communities of Practice: a Guide to Managing Knowledge.* Boston: Harvard Business School Press.

Whalen, M. R., and Zimmerman, D. H. (1990). Describing trouble: practical epistemology in citizen calls to the police 1. *Lang Soc.* 19, 465–492. doi: 10.1017/S0047404500014779

# Instructional Design for Accelerated Macrocognitive Expertise in the Baseball Workplace

Peter J. Fadde *

Department of Curriculum and Instruction, Southern Illinois University, Carbondale, IL, USA

The goal of accelerating expertise can leave researchers and trainers in human factors, naturalistic decision making, sport science, and expertise studies concerned about seemingly insufficient application of expert performance theories, findings and methods for training macrocognitive aspects of human performance. Video-occlusion methods perfected by sports expertise researchers have great instructional utility, in some cases offering an effective and inexpensive alternative to high-fidelity simulation. A key problem for instructional designers seems to be that expertise research done in laboratory and field settings doesn't get adequately translated into workplace training. Therefore, this article presents a framework for better linkage of expertise research/training across laboratory, field, and workplace settings. It also uses a case study to trace the development and implementation of a macrocognitive training program in the very challenging workplace of the baseball batters' box. This training, which was embedded for a full season in a college baseball team, targeted the perceptual-cognitive skill of pitch recognition that allows expert batters to circumvent limitations of human reaction time in order to hit a 90 mile-per-hour slider. While baseball batting has few analogous skills outside of sports, the instructional design principles of the training program developed to improve batting have wider applicability and implications. Its core operational principle, supported by information processing models but challenged by ecological models, decouples the perception-action link for targeted part-task training of the perception component, in much the same way that motor components routinely are isolated to leverage instructional efficiencies. After targeted perceptual training, perception and action were recoupled via transfer-appropriate tasks inspired by *in situ* research tasks. Using NCAA published statistics as performance measures, the cooperating team improved from middling performance to first in their conference in Runs Scored and team Batting Average. This case suggests that, beyond the usual considerations of effectiveness and efficiency, there are four challenges to embedded training in the workplace setting —namely: duration, curriculum, limited resources, and buy in. In the case reported here, and potentially in many domains beyond sports, part-task perceptual-cognitive training can improve targeted macrocognitive skills and thereby improve full-skill performance.

Keywords: perceptual-cognitive, pitch recognition, baseball, macrocognition, expertise

# INTRODUCTION

Sport has long been considered a productive test bed for research on expert performance and training that can potentially accelerate the expertise of performers in military domains (Ward et al., 2008), and other contexts that require macrocognition (defined as cognitive adjustments to performance complexity, *cf*. Klein, 2010). Macrocognitive skills such as anticipation and rapid decision making (Eccles et al., 2008) can potentially be accelerated using *expertise-based training* (XBT) that draws upon the theories, findings, and methods of expertise research in order to design training programs that can efficiently and effectively train expertise in workplace settings (Fadde, 2009a, 2013). XBT focuses on part-task training of cognitive subskills, such as the recognition component of Klein's (1998) model of recognition-primed decision making (Fadde, 2009b). XBT was largely developed in the realm of high-performance sports but also has been applied to accelerating expertise in domains as disparate as classroom teaching (Fadde and Sullivan, 2013), online masters' programs (Tokmak et al., 2013), nursing education (Razer et al., 2015), and peer academic advising (Blair, 2015).

For readers who understand and appreciate expert performance in baseball (or perhaps cricket), this case study provides a deep dive into the pitcher-batter matchup that is at the heart of the sport. For others, the primary points of interest relate to designing training programs that not only apply expert performance research to the task of accelerating expertise but also present research opportunities. Importantly, research design takes a distinctly secondary role to workplace constraints in training-based research. What training-based research projects can offer to the expertise research community are, first, satisfaction with successful implementation of research and, second, insights from fit-in-field modifications that can suggest new basic research questions.
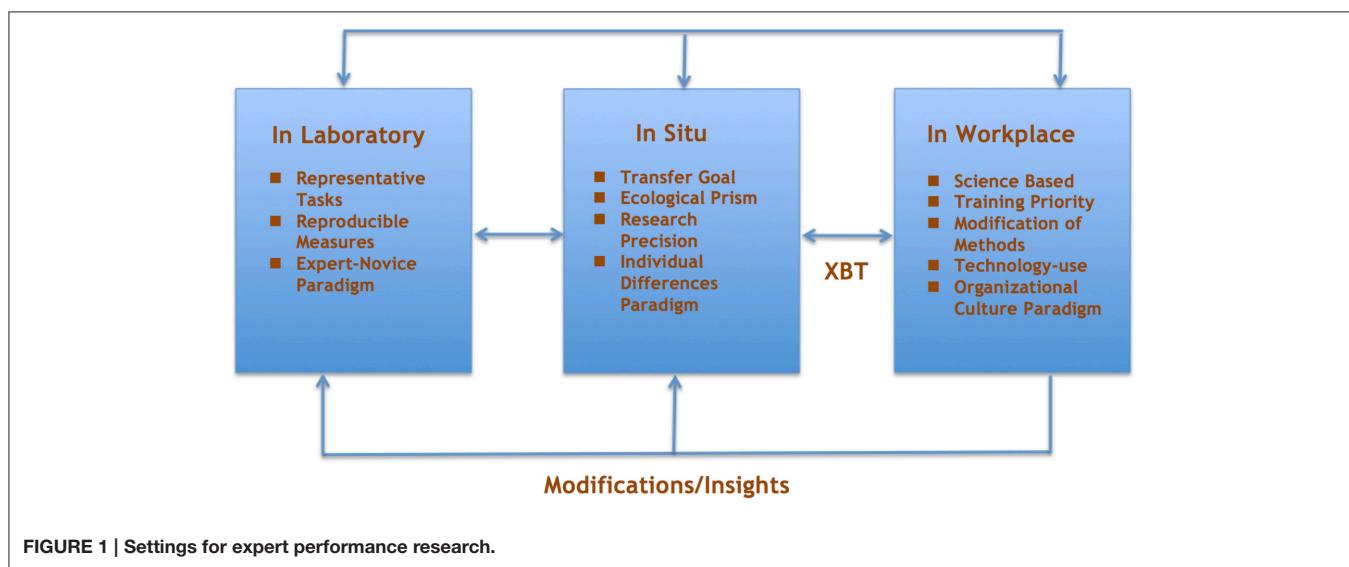
The three settings for expertise research and training shown in **Figure 1** are adapted from a three-stage expert performance model proposed by Williams and Ericsson (2005). Replacing stages with settings in the model emphasizes continuing and iterative processes rather than linear relationships.

"Ultimately, if the expert performance approach has validity, it should be demonstrable through the development of skill-sensitive training... to high levels of performance more quickly" (Charness and Tuffiash, 2008, p. 427). This article argues for and demonstrates an approach that, in the hands of professional instructional designers, military trainers, corporate designers-by-assignment, or human factors engineers, makes connections between expert performance research and expertise-based training. The case study then demonstrates adapting expert performance models and methods to training the perceptual-cognitive skill of *pitch recognition* that underlies one of the most extreme of human performances, hitting a pitched baseball traveling at speeds over 90 miles-per-hour and moving in unexpected directions.

## Transitioning Expertise Research to Expertise Training

Chief among the expertise research methods that have been successfully repurposed for expertise-based training is *temporal occlusion* in which subjects are shown film or video clips depicting a participant's view of an opponent, such as a baseball pitcher, cricket bowler, or tennis server. The film or video image is edited to black (occluded) at various points in the opponent's motion or ensuing ball flight. The representative task given to subjects or trainees is to identify the type of pitch or serve and sometimes predict where the ball will end up in the striking zone of a receiving player. The subject or trainee may respond verbally, by ticking an answer sheet, or even by making a realistic motion such as stepping to her backhand or forehand side to indicate serve location (Williams and Grant, 1999). Though occlusion points may vary across studies and sports, researchers, and trainers in this area agree that athletes can train their perceptual abilities by subtracting visual information during training. Most use video as the medium they rely on for training perceptual skills.



**FIGURE 1 | Settings for expert performance research.**

Not every researcher agrees that performance can be decoupled into smaller, more trainable cognitive units. Advocates of ecological dynamics (Davids et al., 2013) and direct perception (Bootsma and Harvey, 1997) argue strongly that decoupling the perception-action link in ballistic striking skills changes the behavior so that it can't be considered a truly representative task. Indeed, an entirely different visual response system, the dorsal stream, seems to be involved when a perception is intertwined with action rather than when perception is separated from action and therefore engages the ventral stream (Farrow and Abernethy, 2003). Distinct camps represent *predictive control* that holds a cognitive-information processing view supporting a pre-action perception stage, and a *prospective control* view based on Gibson's ecological approach to perception that loathes taking perception out of the context of actor and environment (Gray, 2009).

An information processing-based model has more utility from an instructional design perspective because it supports decoupling of the perception-action link for isolated and efficient training. Part-task training is generally more efficient than whole-task training strategies (for example, immersive simulations) that are supported by ecological views. While part-task training makes sense to baseball coaches who have long trained the mechanical components of batting in part-task ways, development of perceptual-cognitive or macrocognitive skills often is assumed to come only with substantial and varied authentic or simulated experience.

Temporal occlusion as a part-task perceptual training method in sports dates to Haskins' (1965) study that trained intermediate tennis players to recognize opponents' ground strokes. Although it predates articulation of the expert performance approach, Haskins' project shows how long the bones of occlusion training have been in place. As an *in situ* pre-test she filmed subjects returning groundstrokes from an opponent and counted frames of film between the opponent contacting the ball and the subject contacting the ball as a measure of response time. After multiple film-occlusion training sessions, subjects (college students) returned to the court and demonstrated a statistically significant improvement in response times. Haskins had not only created an occlusion-based training task but also devised an *in situ* pre/post-test that was ecologically valid for testing transfer of training gains to performance.

In the sport of baseball, Burroughs (1984) used video-occlusion to train pitch recognition as the perceptual-cognitive component of batting. Burroughs also devised an *in situ* occlusion device to test transfer of laboratory-based learning, which will be discussed further in the baseball training case study. *In situ* tasks not only are used to validate video-occlusion methods but also are used to study the relationship of perception and coordinated motor actions (e.g., Abernethy, 1984; Müller and Abernethy, 2006, 2012, 2014). For training purposes, adding *in situ* tasks may make up for the lack of ecological validity in typical video-occlusion laboratory tasks while also leveraging the precision and efficiency of tasks designed to reveal and measure the perceptual skills that underlie the extraordinarily rapid decision making of skilled athletes in many fast-action sports (Williams and Ward, 2003). Expert-novice studies typically do four things to reveal sources of expert advantage:

(1) Identify critical perceptual-cognitive subskills of performance.
(2) Devise representative tasks that target identified subskills and that are repeatable and readily measurable.
(3) Test and compare highly skilled and less-skilled performers (expert-novice paradigm) to verify that performance on the task differentiates skill groups.
(4) Iteratively add and subtract perceptual information or processing time to locate the boundaries of expert advantage.

Expert-novice studies reveal perceptual-cognitive skills that are critical to expert performance, and they also calibrate the representative tasks and methods used. These precisely defined testing tasks can become extremely efficient and effective training tasks, especially when presented in drill-and-practice format with immediate feedback and progressive difficulty (Alessi and Trollip, 2000). The progression from testing expert advantage to training expert advantage can be viewed in the context of baseball, particularly in the performance skill of batting and its perceptual-cognitive subskill of pitch recognition. In a model expert-novice study, Paull and Glencross (1997) compared the performance of more-skilled and less-skilled Australian professional baseball players on a video-occlusion task that involved identifying the type of pitch (fastball or curveball) being thrown by video pitchers. Pitches were occluded at a variety of points before, at, and after the moment-of-release of the pitch. Paull and Glencross identified which occlusion conditions were most predictive of expert-novice differences and Fadde (2006) used these occlusion points in generating video-occlusion items for a pitch recognition training project. Fadde also added instructional design value by creating Pitch Type, Pitch Location (Known Type), Pitch Location (Unknown Type), and Zone Hitting drills. Drills were edited onto separate videotapes, which were segmented by pitcher and occlusion condition.

As shown in **Figure 2**, a researcher/trainer conducting video-occlusion training would select a drill video, play a video pitch, record the player/trainee's verbal input (e.g., "Fastball" or "Strike"), provide immediate and corrective verbal feedback, and play the next video pitch. After completing all of the pitches in a drill, the researcher/trainer told the player his score on the drill. The player could choose to continue with the same drill video, viewing a different pitcher. The player could also view the same pitcher but at a more difficult occlusion point or choose a different video drill.

## Research-Based Training of Pitch Recognition

Video-occlusion tasks presented in a drill-and-practice instructional format have also been programmed into a sophisticated computer-based pitch recognition training application (Axon Sports, 2015). The *Axon Sports* computer program increases the fidelity level of video-occlusion training by using a 65-inch touch-screen video monitor for display. However, the *Axon Sports* program maintains the part-task recognition-only training approach rather than opting to simulate the whole skill of baseball batting, as a recently released virtual reality baseball training app does (Turner, 2015).

FIGURE 2 | Video-Occlusion in lab (1999).

In large part because baseball batting performance in competitive leagues is represented by an array of statistics, researchers have been able to measure effects of pitch recognition training on performance. For example, the *Axon Sports* pitch recognition training application was made available to an NCAA Division-I college baseball team for self-directed use by players during the 2013 baseball season (Belling and Ward, 2015). Effects of the training program were measured by comparing the cooperating team's batting statistics in the season previous to using the *Axon Sports* system with statistics from the 2013 season. Use of the computer software was not guided or tracked by researchers, but was determined to have been effective because of statistically significant increases in the team's home runs, runs scored, and slugging percentage.

Fadde (2006) also demonstrated effects on batting performance associated with video-occlusion training of pitch recognition by comparing the batting performance of a group of players who received training with a control group of players on the cooperating team who did not receive video-occlusion training. Treatment and control groups were compared by ranking batters on the statistics of Batting Average, On-base Percentage, and Slugging Percentage. Using the Mann–Whitney $U$-test, batters in the treatment group ranked higher on all three batting statistics and significantly higher on Batting Average (Fadde, 2006). Despite the demonstrated effects of pitch recognition training on batting performance, however, these methods have yet to be widely adopted by teams as a routine part of preparing high-performance batters (Belling and Ward, 2015).

Not only does limited application of proven perceptual-cognitive training methods limit potential improvement of high-performance athletes but it also limits the potential that many expertise researchers envision for applying the theories, findings, and methods of sports expertise research to the training of macrocognitive skills in domains such as military and law enforcement (Eccles et al., 2008; Ward et al., 2008). Although years of controlled experimental studies have evidenced the expert performance approach (Abernethy, 1999; Williams and Grant, 1999) the expert performance approach is mostly likely to be adopted for training when it meets the instructional design challenge of fitting into existing workplace routines.

## Challenges for Instructional Designers who are Designing Training

When working on macrocognitive skill training, instructional designers need to balance research, training needs, and workplace constraints as they structure training curricula that aim to improve performance skills in the workplace (Richey et al., 2011). There are at least four challenges for an embedded expertise training program:

### Instructional Design Challenge # 1: Duration

What training duration is needed to make a meaningful difference in performance? Since instructional designers prize efficiency, duration is of considerable importance, as are timing and frequency of training events. Most of the perceptual-cognitive training studies reported in sport science literature were experimental training programs of limited duration, often with novice or intermediate trainees. These studies have served to validate perceptual training techniques and technologies but there is no indication that they have been sustained beyond the experimental context. Ideally, training for advanced performers in the workplace should be available when it is needed and individualized to address gaps between desired and delivered performance (Richey et al., 2011).

### Instructional Design Challenge # 2: Curriculum

Does the training program target specific macrocognitive skills associated with expert performance? Are there existing expert-novice academic studies that suggest target skills? If not, is it worth conducting a small-scale study to discover or confirm macrocognitive skills that differentiate known expert performers from less skilled performers, as Blair (2015) did to inform her design of a training program for peer academic coaches? Once target skills are identified then training tasks can be derived from or inspired by the representative tasks used in expert-novice research. Typically representative tasks focus on situation awareness or pattern recognition and involve: (1) Recall, (2) Detection, (3) Categorization, or (4) Prediction (Chi, 2006).

### Instructional Design Challenge # 3: Resource optimization

Can the program be implemented with limited resources? In part because of relatively limited budgets sport expertise researchers have developed approaches such as video-occlusion, which offers high functional fidelity but low psychological fidelity (it doesn't feel real) by decoupling perception and action for efficient and budget-friendly part-task training. Key concerns are if, when and how performers can recouple the perception-action link for transfer from the part-task training to whole-task performance (Farrow, 2013). *In situ* tasks that researchers have devised to measure learning gains can be repurposed as training tasks that enhance ecological validity. A training program implemented with competing athletic teams or other working professionals could include both highly targeted and efficient video-occlusion tasks and also transfer-appropriate *in situ* tasks.

### Instructional Design Challenge # 4: Buy in

Does the program have commitment from the on-the-ground personnel who influence the effort of trainees? Long-term

sustainability often is tied to the initial buy. In sports training access to high-performance athletes, even when researchers are able to attain it (e.g., Hopwood et al., 2011; Mann et al., 2013), is not enough to ensure success. The attitudes of coaches or superiors toward a training program impact how trainees approach its implementation. Buy in cascades through the curricular design. Because baseball is very routinized in its approach to when and where athletes practice certain skills, a training curriculum has to strive to weave its activities into pre-established routines. Minimizing disruption of habits maximizes the chances of true buy in.

## METHODS

### Training Methods: Case Study of Training Baseball Pitch Recognition

The baseball training project reported here embodies the XBT approach that applies, but also modifies, techniques and technologies of expertise research in order to train key perceptual-cognitive skills and thereby accelerate expertise in already skilled performers. This case study with an NCAA Division-I college baseball team in the U. S. would be labeled a holistic design by Yin (2014); Campbell and Stanley (1963) would call it a one-shot case study. The training-based case study was conducted over a 10-month period in 2013–2014. The goal of the project was to create a training program that was based on research but also fit into established practice routines of the cooperating team.

#### Baseball Context: The Pitcher-Batter Matchup

For readers who may not be familiar with baseball a primer is provided (see Appendix) that provides some basic context. The central action of the game, sometimes called the game within the game, is the individual matchup of pitcher and batter. The act of hitting a round baseball with a round bat, which Ted Williams famously called "the single most difficult thing to do in sport" (Williams and Underwood, 1971, p. 3), affords batters very little margin for error in striking a pitched ball squarely and not popping up or grounding out because of off-centered contact.

At high levels of competition, with many pitchers throwing the ball over 90 miles per hour, batters have less than one-half second from release of the pitch until its arrival in the hitting zone (Bahill and LaRitz, 1984). Most batters take about 250 ms to swing a bat, leaving less than 250 ms (literally the blink of an eye) to decide whether to swing at a pitch and, if so, where to direct the swing. Batters can make fine mid-swing adjustments in the timing and direction of their swing, but only within a limited temporal and spatial window. Therefore, a batter's ability to perceive cues—whether consciously or not—from the pitcher's motion, the release of the pitch, and early ball flight can afford batter precious milliseconds of decision time.

#### Buy In: Initiating the Pitch Recognition Training Program

The relationships of the cooperating teams' coaches with each other as well as with the players were central to implementing an innovative training program. The coaching staff (head coach,

hitting coach, pitching coach, and volunteer assistant coach) was entering a second season with the team in 2014. Before the 2013 season the head coach, who had played for the same college and also played several years of minor league baseball, was hired to replace the previous coach. The head coach hired assistant coaches, who started as an intact staff in 2013. The first season with the team consisted of establishing expectations, policies, and procedures. The team had modest success in 2013, finishing in sixth place in their 11-team conference and thereby being the last team eligible for conference's post-season tournament.

After the initial season's experience, the hitting coach felt empowered to express his opinion that the team's top priority preparing for the 2014 season was to improve batters' pitch recognition. The head coach accepted the hitting coach's arguments that improved pitch recognition would lead to better plate discipline (batters refraining from swinging at pitches out of the strike zone), which—in theory—would reduce strike outs, increase bases-on-balls and on-base percentage (a combination of walks and hits), and runs scored per game. The head coach gave the hitting coach authority (although no budget) to design and install a pitch recognition training program. The hitting coach contacted the researcher and asked for help designing an extensive pitch recognition training program. The coach and the researcher undertook the project understanding that it would be developed iteratively since pitch recognition training studies (Burroughs, 1984; Fadde, 2006; Belling and Ward, 2015) used for guidance were limited in duration and integration.

The pitch recognition training program was initiated in September of 2013. All 18 position players (non-pitchers) provided informed consent and volunteered to participate in the pitch recognition training program. At the team's season orientation meeting the researcher gave a presentation on the sport science research behind the occlusion method of training pitch recognition. The head coach affirmed his support of the program and the hitting coach handed out a Hitting Manual that he had written and printed, which included descriptions of the pitch recognition drills.

#### Pitch Recognition Curriculum

Embedded training programs, in comparison to limited duration experimental training programs, need to have a guiding curriculum. While several sport science studies involved fairly sophisticated experimental training programs that included video-occlusion (e.g., Fadde, 2006; Hopwood et al., 2011) they were still limited duration experimental programs. The best example of a curriculum approach was a visual skills program conducted with a college baseball team over the course of 3 years (Clark et al., 2012). The program had distinct pre-season and in-season phases that included several different visual skills techniques and technologies, such as Nike Strobe goggles and Dynavision hand-eye reaction trainer.

For the pitch recognition training program reported here, the hitting coach and the researcher negotiated two key principles: (1) apply the relevant sport science with as much fidelity as reasonably possible, and (2) integrate pitch recognition training into established team practice routines. The later was important

for sustainability of the training approach and was also necessary because of rules enforced by the National Collegiate Athletic Association (NCAA)—the ruling body of U. S. college sports—that restrict the number of direct contact hours per week between coaches and players.

The pitch recognition training program had several phases that made up a curriculum plan:

(1) Fall Practice (mid-August through late October). Players worked with coaches in groups of four for 1 h 2 days a week. The hitting coach worked with players in the outdoor batting cage (a netted area approximately 20 feet by 60 feet) or the indoor batting cage depending on weather conditions. The session consisted of part-task batting drills, some of which were adapted to include elements of pitch recognition.

(2) Winter Workouts (late October through December). Players self-monitored weight room workouts during a non-contact period with coaches, as defined by NCAA regulations. During the non-contact period a laptop computer with a prototype version of the *Axon Sports* computer application was set up in the baseball office and players used it voluntarily. The laptop was available in November and December.

(3) Spring Practice (January through late February). The full team worked with coaches at the practice field, which included batting cage and bullpen area where pitchers practiced pitching to catchers at full pitching distance. Batters did familiar batting drills that had been introduced in fall practice and Bullpen Stand-In drill was added to practice session.

(4) In Season (late February through May). During the season formal practice sessions were limited but some batters chose to incorporate batting cage drills and Bullpen Stand-In into their pre-game preparation. As with weight lifting, the established routine was heavy work during the off season and light maintenance work during the season, which typically included three games per week with travel to around half of the games.

While there have been several pitch recognition training interventions with college baseball teams they either focused on a short time frame (Burroughs, 1984; Fadde, 2006) or made a training technology available to a team for a full season but did not specify instructional activities (Belling and Ward, 2015). This was the first pitch recognition training program that featured a curriculum throughout the in-season and off-season phases of a sports year.

## Computer-Based Pitch Recognition Component

*Axon Sports* provided the cooperating team with a prototype version of their pitch recognition application that ran on a 17-inch touch screen laptop computer (see **Figure 3**). The computer was available in the baseball office for voluntary and self-directed use by players. A player using the computer system would log in and then use menus to build a drill. The player selected:

(1) Pitcher, from three pitchers that had different repertoires of pitches.
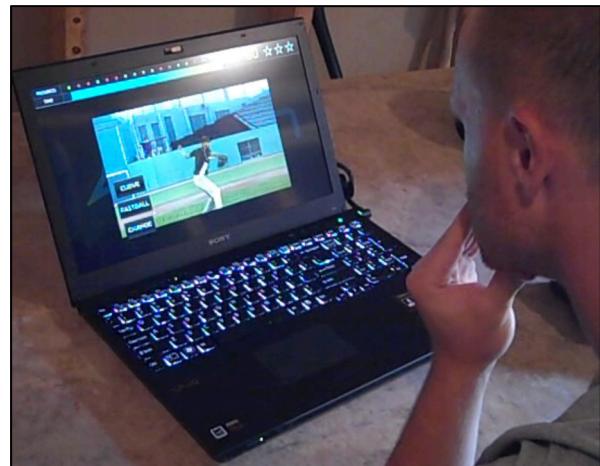


FIGURE 3 | Video-Simulation (courtesy *Axon Sports*).

(2) Batting side, either left-handed or right-handed batter viewpoints.
(3) Drill type, including Pitch Type, Pitch Location, and Zone Hitting.

Players could return to in-progress drills at later sessions. The level of difficulty, which was determined by the amount of ball flight before occlusion, always started at the easiest level and advanced to more difficult levels as players achieved mastery scores, essentially beating the level in video game fashion.

Each round of a drill presented 20 pitches selected from a larger item pool. Players input multiple-choice answers (e.g., Fastball/Curveball/Changeup) by pressing a button on the touch screen. The computer program accepted the player's input, judged correctness of the input, displayed the correct answer, and played an audio tone to indicate correct or incorrect input. The program automatically played the next video pitch and presented a score at the end of the drill. Most drills took about 5 min to complete. At the end of drill the computer would automatically progress the player the next level of the drill if the player had reached criterion score. Although players usage was not tracked, 14 out of 18 players reported that they used the *Axon Sports* computer application at least once and 10 of the players reported that they reached the highest level of progressive difficulty in several video drills.

## Design and Implementation of in situ Batting Cage Drills

The researcher worked with the hitting coach to overlay a pitch recognition element onto several routine batting cage drills that players did during small group workouts. A key challenge was to devise live visual occlusion tasks. Many sport science studies have used liquid crystal occlusion glasses for *in situ* occlusion tasks. Occlusion glasses instantly change from clear to opaque when sent an electronic signal, effectively cutting off the wearer's vision.

Several studies have used occlusion glasses in cricket and baseball batting tasks (e.g., Müller and Abernethy, 2006; Müller et al., 2010, 2015b). In these *in situ* tasks, batters faced a live

baseball pitcher or cricket bowler. In some studies, batters were directed to swing at the pitched ball, even after their vision had been occluded. Researchers gained at least two benefits from *in situ* batting tasks with occlusion glasses. According to ecological dynamics theory (Davids et al., 2013) a batter producing the realistic motor action of swinging his bat should engage the appropriate dorsal stream and maintain the perception-action link. In addition, some studies paired *in situ* occlusion with chronometric analysis using high-speed video cameras and force plates to ascertain precisely when and how a cricket or baseball batter synchs his swing to the movements of the pitcher (Müller et al., 2009, 2014).

While delivering substantial research benefits, however, there are many issues involved with live occlusion tasks. It can take up to 2 h to conduct a test on each subject, which may be tolerated for a one-time experiment but not for routine practice sessions. There is also the possibility of cricket or baseball batters being hit by a pitch when their vision is occluded. Although injury potential can be lessened by using low-impact balls and outfitting batters with elbow guards, the chances of getting hit by a pitched or bowled ball are much higher in training situations than in testing situations because many more pitches are faced in less controlled contexts.

Using live pitchers for *in situ* testing is problematic because the same pitchers can't pitch to every batter. Müller et al. (2015a) argue that the skill of pitch recognition is assumed to generalize across numerous pitchers, so variety is desirable. While certainly a legitimate point for training, testing of pitch recognition that will be used to compare players should certainly be tested against consistent pitchers. The problem can be lessened by using a video pitching/bowling machine, such as *ProBatter*, which displays a video image of a pitcher or bowler that matches the type of ball being delivered (see **Figure 4**). However, the $40-50,000 price of professional grade *ProBatter* is out of the range of most teams.

**Figure 5** shows a pair of liquid crystal display glasses and **Figure 6** shows the patent drawing of a novel device that Burroughs (1984) invented to test and train pitch recognition. The Visual Interruption Systems featured a batting helmet equipped with a plate that would drop in front of a batter's eyes to occlude his vision. The V.I.S. system was triggered by a batter's weight shift while stepping on a force plate.

While occlusion glasses are a valued tools in research settings they may be too complex, expensive, and intrusive to be used in training settings. However, training goals do not require the strict occlusion variations that testing and research goals require. The hitting coach and researcher developed an *in situ* occlusion task that did not require technology but maintained the operational principles (Gibbons, 2009) of occlusion. As shown in **Figure 7**, *Net Occlusion Drill* involved one player standing behind a net drawn across the batting cage and throwing a simulated pitch into the net, effectively occluding ball flight. The player throwing the simulated pitch (usually another batter rather than a real pitcher) showed authentic pitch release cues, such as the skinny wrist many pitchers show when throwing a curveball. The batter read pitch release cues and called the type of pitch aloud. Depending on the objective of the drill (e.g., "hit fastballs") the batter could strengthen the association of recognizing the pitch type and hitting a ball off of the tee.

Net Occlusion Drill has several advantages for the team over batting practice facing a live pitcher. One is that a non-pitcher can throw the stimulus pitches so that pitchers are not being stressed by pitching to batters. Another advantage is that the part-task objective of recognizing pitch types does not become conflated with the full task of hitting the pitch. Net Occlusion Drill represents the second of three levels of video-simulation fidelity proposed by Müller et al. (2015a):

(1) Video Simulation with Non-Motor Response,
(2) Video Simulation or Virtual Reality with Motor Response,
(3) *In-Situ* with Motor Response

### *In situ* in the Bullpen: Attention Occlusion

Another live occlusion drill developed for the pitch recognition training program simulated computer video-occlusion by "standing in" while the team's pitchers were practicing pitching in the bullpen (a designated area at baseball fields where pitchers practice or warm up for a game). A batter would assume his



**FIGURE 5 | Occlusion Glasses (courtesy Translucent Technologies).**



**FIGURE 4 | Video pitching machine (courtesy ProBatter).**
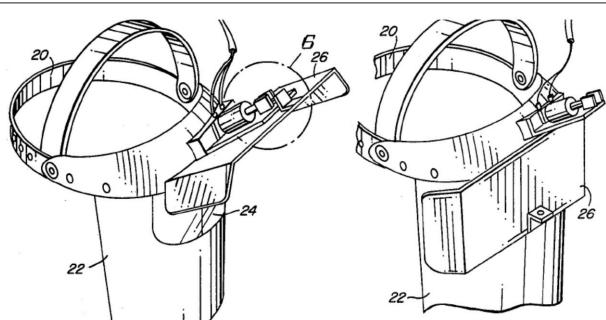


**FIGURE 6 | Visual Interruption System patent illustration.**

**FIGURE 7 | Net Occlusion Drill (https://www.youtube.com/watch?v=ZaXYlcw1uM0).**



**FIGURE 8 | Bullpen Stand-In Pitch Recognition Drill.**

normal position in the batter's box but would not swing his bat (see **Figure 8**). Instead, the batter would call aloud the type of the pitch being delivered before the pitch hit the catcher's mitt. *Bullpen Stand-In Drill* was developed with input from a batting coach who uses it with minor league batters in his major league baseball organization (White, 2014).

Stand-In is a routine practice activity that players have been doing for many years and that is usually associated with tracking pitches from the pitcher's hand to the strike zone. Bullpen Stand-In Drill changes the batter's focus to identifying cues in the pitcher's windup, release of the pitch, and early ball flight. In a video-occlusion context, occlusion removes tracking of pitches by cutting to black during ball flight. In the bullpen, the batter shifts his attention from visual pattern recognition as a System1 cognitive process (Kahneman, 2011) to verbal message construction in System 2, thereby cutting off his attention. Calling out pitch type before the pitch hits the catcher's mitt forces the *attention occlusion* into the time frame from pitch release through 1/3rd ball of flight that expert-novice research found to be the window of maximum expert advantage (Paull and Glencross, 1997). While the cognitive process of attention occlusion is speculative at this point, calling the pitch before it hits the catcher's mitt appears to effectively occlude batters' attention in the critical pitch recognition window.

To be consistent with the computer-based occlusion drills, batters would choose call aloud the pitch type (e.g., Fastball, Curve, or Changeup) or location (ball or strike). Attention occlusion is a level one simulation (Müller et al., 2015a) in that the players' response is verbal rather than a relevant motor movement. When players had been doing Bullpen Stand-In Drill for a couple of weeks the coach gave them a ghost bat that he'd created by sawing off a broken metal bat to about one-foot in length and adding weight to make it feel more like a real bat. The shortened bat meant that the batter could swing at a pitch but without making contact with the ball since bullpens are not designed for batting practice. Allowing batters to swing the ghost bat was satisfying to players and arguably increased the ecological validity (Bootsma and Harvey, 1997) of the Stand-In drill. With the addition of the ghost bat, Bullpen Stand-In Pitch Drill became a level-two simulation in which batters input their pitch recognition verbally and input their swing decision with an authentic movement.

Several of the players were initially reluctant to call pitches out loud, perhaps because it made their mistakes public. The coach countered by reminding players that, "If you're getting them all right, you're doing it all wrong." He wanted players leaving their comfort zone to call pitches earlier. Bullpen Stand-In Drill needed to be carefully monitored to have the desired cognitive training effect. When executed properly, though, it captured value of *in situ* training while addressing several issues associated with occlusion glasses. It did not require expensive or complex technology and it took advantage of real pitchers without adding to their pitching load.

## Research Methods: Procedures used in the Study

Participants in the training program included all 18 of the position players on the cooperating team. The mean age of the participants was 20.7 years. All participants were white males. The participants had been on the cooperating team's roster for an average of 2.5 years at the start of the project. All of the batters who volunteered to participate in the pitch recognition training program received training, so no internal control group

of untrained batters was designated as done in previous studies of pitch recognition training (Fadde, 2006). The unit of analysis was batting performance of the team as a group.

## Batting Performance

The primary research question was whether the embedded pitch recognition training program would lead to improvements in team batting performance. The independent variable was the pitch recognition program in its entirety, including the computer-based video-occlusion application and the *in situ* pitch recognition drills. There were two dependent variables, both based on season team batting statistics published by the NCAA (2015). The first DV was the batting performance of the cooperating team in baseline (2013), implementation (2014), and adoption (2015) seasons. Serving as control, batting statistics were compared to the mean values on the same statistics of all the teams in the cooperating team's athletic conference. The second DV was change in the cooperating team's batting statistics from baseline season (2013) to implementation season (2014) seasons.

Analysis of the change in the team's batting performance was compared to change in batting performance over the same seasons by a comparable team in the same athletic conference. The team designated as the comparison team was the conference team most similar to the cooperating team. Both the cooperating team and the comparison team returned 7 out of 8 batters from their 2013 starting lineups for the 2014 season. Both teams made the 6-team post-season tournament in both the 2013 and 2014 seasons. Both teams improved their win-loss record and position in the conference standings from 2013 to 2014, with the cooperating team winning the 2014 conference regular season championship and the comparison team winning the conference's 2014 post-season tournament. Comparing the cooperating team's change in performance to a selected and comparable conference team, rather than using the mean performance of the whole conference, let the research address the coaches' question of whether any improved performance was "beyond what would be expected from a good team getting better."

The batting statistics analyzed were the team performance measure of Runs-per-Game along with the individual performance measures of Batting Average, On-base percentage (which includes walks and hits), and Slugging Percentage (which counts all bases and is considered to be a measure of power hitting)—three statistics that thought to provide a rounded profile of batting performance (Weinberg, 2014). Other statistics analyzed included Walk Rate, Strikeout Rate, and Walk-to-Strikeout Ratio that are considered to represent plate discipline (Panas, 2010). Scoring (Runs-per-Game) is the most basic measure of team offensive performance; Walk-to-Strikeout Ratio is the most basic measure of individual plate discipline.

## Pitch Recognition Testing

As noted earlier, testing and training have a close relationship in expertise-based training. The pitch recognition training project described here offered several opportunities to test not only for group differences, as has been done in the expert-novice research paradigm, but also test for individual differences and individual

development as sport science researchers are just beginning to pursue. As a training project, however, ideal testing conditions for research purposes were sometimes compromised for the sake of team preparation and competition.

After the pitch recognition training program was underway, a validated video-occlusion Pitch Recognition (PR) test became available and was administered to batters on the cooperating team. Later, a second video-occlusion pitch recognition test became available and was also administered to the cooperating team. Both PR video tests showed pitches from a perspective closely, but not exactly, depicting the view of a participating batter. However, the tests differed in the occlusion points that were used. While the seminal laboratory-based expert-novice study of pitch recognition (Paull and Glencross, 1997) used an array of occlusion points cutting off pitches before, at, and after the pitcher released the pitch, testing professional baseball batters in the field required researchers to construct shorter video-occlusion tests.

The first video occlusion test developed for testing professional players, heretofore called the Pre-Release Test, used video clips of pitches that were occluded at Release of the pitch and at two occlusion points before Release. The Pre-Release video-occlusion test was formally validated and used to test professional players competing in the Australian Baseball League (Moore and Müller, 2014) and later used to test minor league players in the United States (Müller and Fadde, 2016). The second test, heretofore called the Post-Release Test, was developed later and featured pitches that were occluded at Release and at two occlusion points after the pitcher released the pitch. Both the Pre-Release and Post-Release tests were administered to batters on the cooperating team, which allowed several questions about pitch recognition testing to be addressed:

(1) Would either or both PR tests differentiate groups of batters by skill level?
(2) Would either or both PR test correlate with batting performance?
(3) Would the Pre-Release and Post-Release tests correlate with each other?
(4) What insights might be gained from PR testing for coaching purposes?

The Pre-Release PR test was administered in the fall of 2014. The 2014 baseball season finished in May and the test was administered at the beginning of the next school year (2014–2015), which is considered to be part of the 2015 season. The college baseball season is split into a fall period with organized practice and the competition portion of the season in the spring of the next calendar year. Of 20 players who took the PR test in Fall 2014, 10 played regularly (100+ Plate Appearances) in the 2014 season or would be regular players in the 2015 season. The other 10 players played part-time. The PR scores of these two groups were compared in an adaptation of expert-novice methodology. Batters' individual scores on the Pre-Release PR test were also correlated with season batting statistics of seven batters who had been regular starting players in the 2014 season.

The Post-Release PR test was administered twice in the fall of 2015, about 6 weeks apart. Scores on the first Post-Release PR test were correlated with scores on the Pre-Release PR test. The two administrations of the Post-Release PR test were correlated with each other to address the question of whether pitch recognition is a stable trait of batters or a fluctuating state. Since regular and extensive testing of the pitch recognition skill of batters in many contexts will require different PR tests it is important to develop methods of validating new tests. The coaches of the cooperating team embraced testing for development of their players as well as advancing the science of pitch recognition testing.

## RESULTS

### Team Batting Performance

**Table 1** shows the cooperating team's batting statistics for the 2013 season (baseline), the 2014 season (implementation), and the 2015 season (adoption). The mean batting statistics of all the teams in the cooperating team's athletic conference serve as control. Change in the statistics of *Runs-per-Game* and *Walk-to-Strikeout Ratio* (BB/K) are bolded in **Tables 1**, **2** because these are the most relevant statistical representations of team offense and individual plate discipline, which is defined as swinging at pitches that are in the strike zone and refraining from swinging at pitches that are out of the strike zone. Values that are shown in parentheses in the Differences columns indicate lower performance by the cooperating team. Strikeouts are reverse scored; a lower number is considered to be a better performance.

In the 2013 season, the cooperating team was below conference mean on almost all batting statistics. In the 2014 season, which included pitch recognition training, the cooperating team was higher than conference means on all of the analyzed batting statistics. In 2015 the cooperating team's batting statistics were again consistently better than the mean scores of the conference, with the exception of strikeouts. As context in interpreting batting statistics, general benchmarks at the major league level include: 0.300 for Batting Average, 0.375 for On-base Percentage, 0.450 for Slugging Percentage, and 0.500 for Walk-to-Strikeout Ratio (BB/K).

While **Table 1** shows clearly superior batting performance in the implementation year (2014) following pitch recognition

training, the central question of whether pitch recognition training was associated with improvement in batting performance from baseline to implementation seasons was evaluated by comparison with improved performance of a similarly successful team (see **Table 2**).

An effect size for improvement in the key batting statistic of Walk-to-Strikeout Ratio (BB/K) from 2013 to 2014 was calculated using the mean BB/K of 12 batters on the 2013 roster (mean = 0.51; $sd$ = 0.22) and 11 batters, including six hold-overs from 2013, on the 2014 roster (mean = 0.80; $sd$ = 0.37) who had a minimum of 50 plate appearances (as Belling and Ward, 2015). The effect size was large ($d$ = 0.953) and significant ($p$ = 0.017) at $p < 0.05$.

While coaches were satisfied with percentage of change as evidence of improvement, as shown in **Tables 1, 2**, the research question of whether pitch recognition training was associated with improved overall batting performance required determining the statistical significance of overall performance improvement from 2013 to 2014. Overall season-to-season improvement was assessed by comparing the conference ranks on selected batting statistics in 2013 and 2014 of both the cooperating (training) and the comparison (no training) team (see **Figure 9**). Mann–Whitney $U$-test of rank correlation, scaled for small $n$, was used to compare 2013 and 2014 seasons as a whole for each team. With 11 teams competing in the conference, the top rank score was

**TABLE 2 | Changes in Batting Statistics: Cooperating Team vs. Comparison Team.**

|  | Cooperating Team 2013–2014 Change | | | Comparison Team 2013–2014 Change | | |
|---|---|---|---|---|---|---|
| Runs Per Game | 5.8 | 8.6 | **48%** | 6.6 | 6.8 | **3%** |
| Batting Average | 0.286 | 0.326 | 14% | 0.290 | 0.304 | 5% |
| On-base Pct. | 0.372 | 0.407 | 9% | 0.372 | 0.383 | 3% |
| Slugging Pct. | 0.390 | 0.468 | 20% | 0.413 | 0.464 | 12% |
| Home Runs | 11 | 25 | 127% | 17 | 27 | 59% |
| Base-on-Balls | 108 | 140 | 30% | 127 | 124 | (2%) |
| Strikeouts (K) | 217 | 182 | 16% | 189 | 200 | (6%) |
| BB/K Ratio | 0.50 | 0.77 | **54%** | 0.67 | 0.62 | **(7%)** |

*Change in key PR stats bolded.*

**TABLE 1 | Differences in Batting Statistics: Cooperating Team vs. Conference.**

|  | 2013 | | | 2014 | | | 2015 | | |
|---|---|---|---|---|---|---|---|---|---|
|  | Team | Conf. | Diff. | Team | Conf. | Diff. | Team | Conf. | Diff. |
| Runs Per Game | 5.8 | 6.2 | **(6%)** | 8.6 | 6.3 | **37%** | 9.4 | 7.2 | **17%** |
| Batting Average | 0.286 | 0.291 | (2%) | 0.326 | 0.290 | 12% | 0.324 | 0.301 | 8% |
| On-base Pct. | 0.372 | 0.371 | – | 0.407 | 0.375 | 11% | 0.419 | 0.385 | 9% |
| Slugging Pct. | 0.390 | 0.418 | (7%) | 0.468 | 0.422 | 11% | 0.519 | 0.435 | 19% |
| Home Runs | 11 | 21 | (48%) | 25 | 23 | 9% | 39 | 35 | 11% |
| Base-on-Balls | 108 | 111 | (3%) | 140 | 120 | 17% | 171 | 126 | 36% |
| Strikeouts (K) | 217 | 200 | (9%) | 182 | 199 | 9% | 219 | 203 | (8%) |
| BB/K Ratio | 0.50 | 0.56 | **(11%)** | 0.77 | 0.60 | **28%** | 0.78 | 0.62 | **26%** |

*Change in key PR stats bolded.*
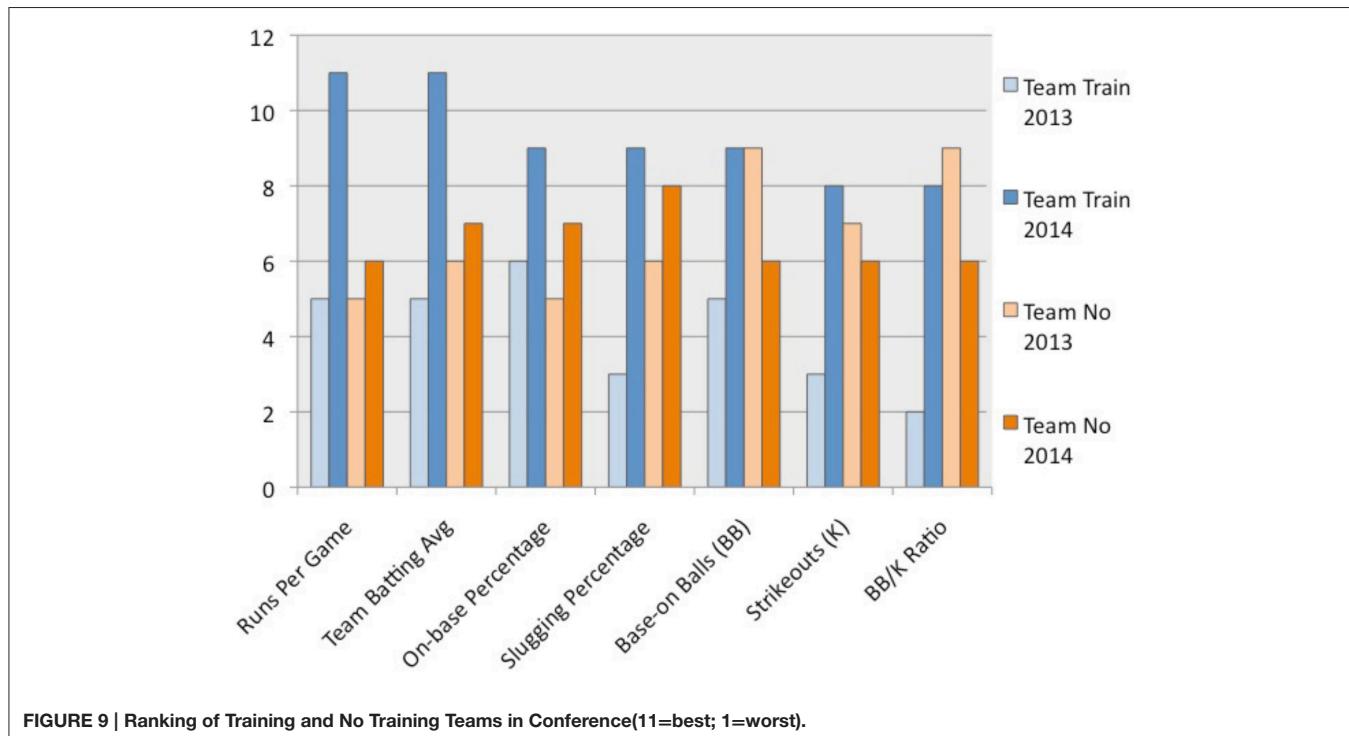
**FIGURE 9 | Ranking of Training and No Training Teams in Conference(11=best; 1=worst).**

**TABLE 3 | Change in Batting Statistics for Cooperating Team.**

|                 | 2013–2014 Change | | | 2014–2015 Change | | |
|-----------------|-------|-------|------|-------|-------|------|
| Runs Per Game   | 5.8   | 8.6   | 48%  | 8.6   | 9.4   | 9%   |
| Batting Average | 0.286 | 0.326 | 14%  | 0.326 | 0.324 | <1%  |
| On-base Pct.    | 0.372 | 0.407 | 9%   | 0.407 | 0.419 | 3%   |
| Slugging Pct.   | 0.390 | 0.468 | 20%  | 0.468 | 0.519 | 11%  |
| Home Runs       | 11    | 25    | 127% | 25    | 39    | 56%  |
| Base-on-Balls   | 108   | 140   | 30%  | 140   | 171   | 22%  |
| Strikeouts (K)  | 217   | 182   | 16%  | 182   | 219   | (20%)|
| BB/K Ratio      | 0.50  | 0.77  | 54%  | 0.77  | 0.78  | 1%   |

"11" and "1" was the bottom rank score. Applying a one-tailed analysis with alpha of $p < 0.05$, the cooperating team's overall ranking on the selected batting statistics was significantly higher ($p = 0.0005$) in 2014 than in the 2013 season. The same analysis conducted on the comparison (no training) team's improvement from 2013 to 2014 was not significant ($p = 0.4364$). **Figure 9** graphically displays the cooperating team improvement "beyond expectations of a good team getting better."

After implementation of the pitch recognition training program for the 2014 season, the cooperating team continued to incorporate pitch recognition training, even without the *Axon Sports* computer application or direct involvement of the researcher. Although not as dramatic as the improvement from the baseline season (2013) to the implementation season (2014), the cooperating team continued to improve in the adoption season of 2015 (see **Table 3**).

Of particular note in the 2015 season was the increase in home runs while also increasing in strikeouts, which reflected

the hitting coach's 2015 hitting theme of being more aggressive at the plate. While batters struck out more, they also walked more, hit more home runs, and scored more runs. The 2015 batting statistics counter the common concern of coaches that training pitch recognition may lead to overly selective, and therefore passive, batters.

## Pitch Recognition Testing

The pitch recognition training project offered numerous opportunities for testing the PR skills of batters on the cooperating team. Administering two different video-occlusion PR tests to batters on the cooperating team permitted several questions related to PR testing to be addressed.

(1) Would either or both PR tests differentiate groups of batters by skill level?

Batters on the cooperating team were tested using a validated video-occlusion Pre-Release PR Recognition test (Moore and Müller, 2014) in advance of the 2015 season. The expert-novice paradigm was adapted, as Moore and Müller (2014) did, to compare a group of higher-skilled batters (players who were regularly in the starting lineup in either 2014 or 2015 seasons) with a group of lesser-skilled batters (non-regulars). The higher-skilled group's mean PR score was 58.8 while the less-skilled group's mean PR score was 52.1; the difference ($p = 0.1304$) was non-significant at $p < 0.05$.

(2) Would batters' PR test scores correlate with batting performance?

The Pre-Release PR test scores of individual batters were correlated with the batters' 2014 season batting statistics for

Batting Average, On-Base Percentage, Slugging Percentage, Walk Rate, Strikeout Rate and Walk-to-Strikeout ratio. Using a minimum participation rate of 100 Plate Appearances (Moore and Müller, 2014), 11 of 18 batters tested qualified for the analysis. Using Pearson product moment coefficient, no significant correlations were found (at $p < 0.05$), a finding that is consistent with a study of minor league batters that found a significant correlation only for Walk Rate at one pre-release occlusion point (Müller and Fadde, 2016). The Post-Release PR test scores of six players who had played in the 2015 season were analyzed but did not significantly correlate with any of the batting statistics, in part because of the small number of players. However, the Post-Release PR scores can potentially be correlated with the 2016 batting statistics that will be generated by up to 18 batters who took the at least one version of the Post-Release test in fall 2015.

(3) Would the PR tests correlate with each other?

Only six batters took both the Pre-Release video PR test (in Fall 2014) and the Post-Release video PR test (in Fall 2015). The correlation between batters' scores on the two tests was moderate to strong ($r = 0.707$) but not significant ($p = 0.117$). The finding suggests that one validated video PR test can potentially be used to validate a second video PR test, but with further investigation needed.

The Post-Release PR test was administered twice in fall 2015 with 14 out of 18 batters on the cooperating team's roster completing both tests, which were given about 6 weeks apart. Mean PR score on the first Post-Release test was 62.7 and the mean score on the second administration of the Post-Release test was 61.2, producing a moderately strong correlation ($r = 0.53$) that approached significance ($p = 0.052$). The correlation suggests that taking the video test with no item or summary feedback leads to minimal, if any, learning effect. At least provisionally, either PR video test could be used as both a pre-test and a post-test for research or training purposes. It also suggests that the PR test measures a fairly stable trait. Whether and to what extent batters' pitch recognition skill can change as a result of training, experience, or maturation—and whether changes can be measured with a video PR test—remain to be investigated. Being able to use the same test for repeated measures can be an important tool for coaches as well as researchers in addressing these questions.

(4) What insights might be gained from PR testing for coaching purposes?

The Pre-Release PR score of batters on the cooperating team was 55.40 ($sd = 11.12$), with PR scores ranging from 33 to 75. By comparison, 34 minor league baseball players completing the same video-occlusion test scored an overall mean PR score of 60.25 (Müller and Fadde, 2016). As noted above, players' individual scores on the Pre-Release PR test did not correlate directly with any individual batting statistics. One reason for the lack of correlation between test scores and performance is that hitting is an exceptionally complex system of psychological, cognitive, perceptual, and psychomotor subskills. Even a highly valid test of any one component skill is unlikely to predict

overall skill performance. However, an astute coach can use measurement data on any or all components to inform selection and development of players.

Two batters' scores on the Pre-Release PR test illustrate how the same PR test score can have different implications for different players. The batters both achieved a score of 75 on the Pre-Release test, the highest scores of the 18 players taking the test. One of the players was a senior backup catcher whose primary hitting attribute was a good eye, that is, the ability to predict which pitches would or would not be in the strike zone. However, limited athleticism and several injuries over his career had led to limited playing time other than pinch hitting (batting in place of another player). His high PR score was consistent with his value and role on the team.

The other player to score a 75 on the Pre-Release PR test illustrates a potential use of PR testing to inform coaches' decisions about playing time or training approaches. The batter had a 0.242 Batting Average (compared to mean team Batting Average of 0.306) as a semi-regular in the 2014 season, his sophomore season. He then enjoyed a breakout season in 2015 with a Batting Average of 0.318 (team mean BA = 0.324) and hit a team-leading 13 home runs. Although providing only anecdotal evidence, the coaches' decision to keep this player in the lineup despite relatively poor batting performance appears to have been affirmed when the player's physical maturity and batting technique caught up with his advanced batting eye. PR testing has considerable utility to coaches if it reveals or confirms that a player has perceptual-cognitive skills that may not effect his progression from competence-to-proficiency but may play a role in accelerating the player's progression to expertise (Dreyfus, 2004).

# DISCUSSION

## Limitations and Future Research

The study focused on just one macrocognitive aspect of baseball batting, pitch recognition, at the expense of other macrocognitive skills such as option generation involved in anticipating types of pitches based on game situations and opponent tendencies (Gray, 2002; Lebiere et al., 2003). Future studies could incorporate what Ted Williams called "proper thinking" about pitch probabilities into pitch recognition training (Ward et al., 2013; Cañal-Bruland and Mann, 2015; Cañal-Bruland et al., 2015).

The validated Pre-Release and Post-Release pitch recognition tests were not yet available when the pitch recognition training project started, so systematic pre/post-testing of PR skills before and after the initial implementation season was not possible. Although the cooperating team plans to continue testing and training pitch recognition it is unlikely that an entire group of players will be tested and start a training program at the same time. When opportunities for embedded training with competing teams arise, researchers must balance the value of critical pre-implementation testing with the need to fit into a team's established routines.

Testing, both video-based (Belling et al., 2015) and *in situ*, should have a larger role in future training programs, in part to address important and largely unknown questions about

the state-vs.-trait nature of macrocognitive skills such as pitch recognition. Are these skills stable traits or can they be improved through targeted training? How much training, and of what type, leads to the most improvement? What are the minimum levels of physical, cognitive, and technical development needed in order to benefit from expertise-based training? Would the training methods used with Division-I college baseball players also work with more advanced professional batters, or with high school or even younger batters? Hopefully, embedded training/research projects address these questions in the process of implementing authentic macrocognitive training programs in military and other time-restricted, high-stress performance domains (Fadde, 2010, 2012).

## CONCLUSIONS

The findings of this case study don't support generalizing results beyond the specific team, training program, and performance domain. However, the performance gains of the cooperating team were dramatic enough to invite other baseball teams to develop, implement, and assess at least a portion of the pitch recognition curriculum developed in this study. In addition, researchers and trainers working in other domains may consider developing, implementing, researching, and reporting training programs similarly focused on specific and known macrocognitive components of performance in a variety of high-performance jobs. In military contexts, for example, recognition-based tasks such as patrol leaders spotting roadside explosive devices or landing signal officers waving off a pilot may be amenable to accelerated expertise through expertise-based training.

XBT champions efficiency, even in the nebulous realm of expertise. By focusing on instructional methods rather than technology-driven delivery systems (Clark, 1983), by holding to operational principles of a learning and performance systems (Gibbons, 2009) rather than satisfying, but not always optimal, whole-task learning experiences we are more likely to avoid building the wrong simulation (Foshay, 2006). Instead, we can target identified macrocognitive subskills of expert performance using representative tasks that favor cognitive fidelity over physical fidelity (Fadde et al., 2007) and thereby accelerate expertise in systematic and affordable ways.

Reflecting on the framework of research in laboratory, field, and workplace settings (see **Figure 1**) the theories, findings, models, methods, and representative tasks that emerge from expertise research deserve to be more widely applied in embedded macrocognitive training programs. In many domains and workplaces much more important than baseball the instructional design, human factors, and expert performance communities need to more quickly get more performers over the bars of expertise and expert performance (Hoffman et al., 2014).

Embedded training programs are likely to have widely varied content, contexts, and fidelity of implementation, but if they focus on key operational principles derived from laboratory and field research settings then they can potentially advance both research-based practice and practice-based research. Ideally, the workarounds and modifications that inevitably emerge

from embedded real-world training programs should feed questions back to the basic research community so that they can be thoroughly investigated in controlled laboratory and field research settings. An example is the *attention occlusion* method that was adapted to contextual constraints but should be validated in the laboratory, perhaps using EEG instrumentation to observe specific points in time where a literal spike of pitch recognition is observed (Houdé et al., 2000; Muraskin et al., 2013; Park et al., 2015).

Instructional design theorists and practitioners have important roles to play in collaborating with cognitive psychologists in the human factors and naturalistic decision making communities to develop approaches that train macrocognition in the workplace (Fadde and Klein, 2012). The theories, findings, and methods of expert performance research need to be translated into focused workplace training programs that meet the challenges of duration, curriculum development, resource optimization, and buy in from on-the-ground practitioners. In summary, expertise-based training that applies research methods such as temporal occlusion in the context of workplace training can provide efficient and effective methods of systematically training aspects of performance that are typically assumed to come only with innate talent or massed experience.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

## FUNDING

## ACKNOWLEDGMENTS

## SUPPLEMENTARY MATERIAL

### Videos

1) "Doing Damage at the Plate by Training Pitch Recognition," an informational clinic about the pitch recognition training program presented at the American Baseball Coaches Association conference, Orlando FL, January 2015: https://www.youtube.com/watch?v=ukuKbj_5sUQ

2) "Look, 'Ma, No Hands," a presentation about applying sports expertise methods to training in military domains. Inter-service/Industry Technology, Simulation, and Education Conference (I/ITSEC), Orlando FL, December 2010: https://vimeo.com/36425360

# REFERENCES

Abernethy, B. (1984). "Skill in cricket batting: laboratory and applied evidence," in *Proceedings of the VII Commonwealth and International Conference on Sport, Physical Education, Recreation and Dance,* Vol. 7. (Brisbane, QLD: University of Queensland), 35–50.

Abernethy, B. (1999). The 1997 Coleman Roberts Griffith address movement expertise: a juncture between psychology theory and practice. *J. Appl. Sport Psychol.* 11, 126–141. doi: 10.1080/10413209908402955

Alessi, S., and Trollip, S. (2000). *Multimedia for Learning: Methods and Development.* Boston, MA: Allyn and Bacon.

Axon Sports (2015). *Training Above the Neck.* Available online at: http://www.axonsports.com/

Bahill, A. T., and LaRitz, T. (1984). Why can't batters keep their eyes on the ball. *Am. Sci.* 72, 249–253.

Belling, P. K., and Ward, P. (2015). "Time to start training: a review of cognitive research in sport and bridging the gap from academia to the field," in *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics, Procedia Manufacturing,* Vol. 3 (Las Vegas, NV), 1219–1224.

Belling, P. K., Sada, J., and Ward, P. (2015). "Assessing hitting skill in baseball using simulated and representative tasks," in *Proceedings of the International Conference on Naturalistic Decision Making 2015.* (McLean, VA).

Blair, L. (2015). "Pedal to the metal: accelerating expertise through mixed methods research," in *Paper Presented at the Association for Educational Communications and Technology Conference.* (Indianapolis, IN).

Bootsma, R. J., and Harvey, L. (1997). Perception and action in sport: half-time comments on the match. *J. Sports Sci.* 15, 641–642.

Burroughs, W. (1984). Visual simulation training of baseball batters. *Int. J. Sport Psychol.* 15, 117–126.

Campbell, D. T., and Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research.* Boston, MA: Houghton Mifflin Co.

Cañal-Bruland, R., and Mann, D. L. (2015). Time to broaden the scope of research on anticipatory behavior: a case for the role of probabilistic information. *Front. Psychol.* 6:1518. doi: 10.3389/fpsyg.2015.01518

Cañal-Bruland, R., Filius, M. A., and Oudejans, R. R. D. (2015). Sitting on a fastball. *J. Motor Behav.* 47, 267–270. doi: 10.1080/00222895.2014.976167

Charness, N., and Tuffiash, M. (2008). The role of expertise research and human factors in capturing, explaining, and producing superior performance. *Hum. Factors* 50, 427–432. doi: 10.1518/001872008X312206

Chi, M. T. H. (2006). "Laboratory methods for assessing experts' and novices' knowledge," in *The Cambridge Handbook of Expertise and Expert Performance,* eds K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman (Cambridge: Cambridge University Press), 167–184.

Clark, J. F., Ellis, J. K., Bench, J., Khoury, J., and Graman, P. (2012). High-performance vision training improves batting statistics for University of Cincinnati baseball players. *PLoS ONE* 7:e29109, 1–6. doi: 10.1371/journal.pone.0029109

Clark, R. E. (1983). Reconsidering research on learning from media. *Rev. Educ. Res.* 53, 445–459. doi: 10.3102/00346543053004445

Davids, K., Araújo, D., Vilar, L., Renshaw, I., and Pinder, R. (2013). An ecological dynamics approach to skill acquisition: implications for development of talent in sport. *Talent Dev. Excell.* 5, 21–34.

Dreyfus, S. E. (2004). The Five-Stage Model of adult skills acquisition. *Bull. Sci. Technol. Soc.* 24, 177–179. doi: 10.1177/0270467604264992

Eccles, D. W., Ward, P., Janelle, C. M., Woodman, T., and LeScanff, C. (2008). "Shared interests in solving common problems: how sport psychology might inform human factors and ergonomics," in *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting* (New York, NY), 743–747.

Fadde, P. J. (2006). Interactive video training of perceptual decision making in the sport of baseball. *Technol. Instr. Cogn. Learn.* 4, 265–285.

Fadde, P. J. (2009a). Expertise-based training: Getting more learners over the bar in less time. *Technol. Instr. Cogn. Learn.* 7, 171–197.

Fadde, P. J. (2009b). Instructional design for advanced learners: training expert recognition skills to hasten expertise. *Educ. Technol. Res. Dev.* 57, 359–376. doi: 10.1007/s11423-007-9046-5

Fadde, P. J. (2010). "Look 'Ma, no hands: part-task training of perceptual-cognitive skills to accelerate psychomotor expertise," in *Proceedings of the Interservice/Industry Technology, Simulation, and Education Conference (I/ITSEC).* (Orlando, FL).

Fadde, P. J. (2012). "What's wrong with this picture? Video-annotation with expert-model feedback as a method of accelerating novices' situation awareness," in *Proceedings of the Interservice/Industry Technology, Simulation, and Education Conference (I/ITSEC).* (Orlando, FL).

Fadde, P. J. (2013). "Accelerating the acquisition of intuitive decision making through expertise-based training (XBT)," in *Proceedings of the Interservice/Industry Technology, Simulation, and Education Conference (I/ITSEC).* (Orlando, FL).

Fadde, P. J., and Klein, G. (2012). Accelerating expertise using action learning activities. *Cogn. Technol.* 17, 11–18.

Fadde, P. J., and Sullivan, P. (2013). "Using interactive video to develop pre-service teachers' classroom awareness,"in *Contemporary Issues in Technology and Teacher Education, Vol. 13.* Available online at: http://www.citejournal.org/vol13/iss2/general/article1.cfm

Fadde, P. J., Sarmiento, R., and Arevalo, J. Jr. (2007). "Low physical fidelity / high cognitive fidelity simulation training of expert recognition skills," in *Proceedings of Society for Applied Learning Technology (SALT) New Learning Technologies Conference.* (Orlando, FL). [CD-ROM].

Farrow, D. (2013). Practice-enhancing technology: a review of perceptual training applications in sport. *Sports Technol.* 6, 170–176. doi: 10.1080/19346182.2013.875031

Farrow, D., and Abernethy, B. (2003). Do expertise and the degree of perception–action coupling affect natural anticipatory performance? *Perception* 32, 1127–1139. doi: 10.1068/p3323

Foshay, W. R. (2006). Building the wrong simulation: matching instructional intent in teaching problem solving to simulation architecture. *Technol. Instr. Cogn. Learn.* 3, 63–72.

Gibbons, A. S. (2009). The value of the operational principle in instructional design. *Educ. Technol.* 49, 3–9.

Gray, R. (2002). Markov at the bat: a model of cognitive processing in baseball batters. *Psychol. Sci.* 13, 543–548. doi: 10.1111/1467-9280.00495

Gray, R. (2009). A model of motor inhibition for a complex skill: baseball hitting. *J. Exp. Psychol.* 15, 91–105. doi: 10.1037/a0015591

Haskins, M. J. (1965). Development of a response-recognition training film in tennis. *Percept. Motor Skills* 21, 207–211. doi: 10.2466/pms.1965.21.1.207

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., and Andrews, D. H. (2014). *Accelerated Expertise: Training for High Proficiency in a Complex World.* New York, NY: Psychology Press.

Hopwood, M. J., Mann, D. L., Farrow, D., and Nielsen, T. (2011). Does visual-perceptual training augment the fielding performance of skilled cricketers? *Int. J. Sports Sci. Coaching* 6, 523–535. doi: 10.1260/1747-9541.6.4.523

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., et al. (2000). Shifting from the perceptual brain to the logical brain: the neural impact of cognitive inhibition training. *J. Cogn. Neurosci.* 12, 721–728. doi: 10.1162/089892900562525

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York, NY: Farrar, Stauss and Giroux.

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

Klein, G. (2010). "Macrocognitive measures for evaluating cognitive work," in *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams,* eds E. S. Patterson and J. E. Miller (Farnham: Ashgate), 47–66.

Lebiere, C., Gray, R., Salvucci, D., and West, R. L. (2003). "Choice and learning under uncertainty: a case study in baseball batting," in *Proceedings of the 25th Annual Conference of the Cognitive Science Society.* (Mahwah, NJ: Lawrence Erlbaum), 704–709.

Mann, D. L., Spratford, W., and Abernethy, B. (2013). The head tracks and gaze predicts: how the world's best batters hit a ball. *PLoS ONE* 8:e58289. doi: 10.1371/journal.pone.0058289

Moore, C. G., and Müller, S. (2014). Transfer of expert visual anticipation to a similar domain. *Q. J. Exp. Psychol.* 67, 186–196. doi: 10.1080/17470218.2013.798003

Müller, S., Abernethy, B., Eid, M., McBean, R., and Rose, M. (2010). Expertise and the spatio-temporal characteristics of anticipatory information pick-up from complex movement patterns. *Perception* 39, 745–760. doi: 10.1068/p6438

Müller, S., Abernethy, B., Reece, J., Rose, M., Eid, M., McBean, R. et al. (2009). An *in-situ* examination of the timing of information pick-up for interception by cricket batsmen of different skill levels. *Psychol. Sport Exerc.* 10, 644–652. doi: 10.1016/j.psychsport.2009.04.002

Müller, S., and Abernethy, B. (2006). Batting with occluded vision: an *in-situ* examination of the information pick-up and interceptive skills of high and low skilled cricket batsmen. *J. Sci. Med. Sport* 9, 446–458. doi: 10.1016/j.jsams.2006.03.029

Müller, S., and Abernethy, B. (2012). Expert anticipatory skill in striking sports: a review and a model. *Res. Q. Exerc. Sport* 83, 175–187. doi: 10.1080/02701367.2012.10599848

Müller, S., and Abernethy, B. (2014). An expertise approach to training anticipation using temporal occlusion in a natural setting. *Technol. Instr. Cogn. Learn.* 9, 295–312.

Müller, S., and Fadde, P. J. (2016). The relationship between visual anticipation and baseball batting game statistics. *J. Appl. Sport Psychol.* 28, 49–61. doi: 10.1080/10413200.2015.1058867

Müller, S., Brenton, J., and Rosalie, S. M. (2015a). Methodological considerations for investigating expert interceptive skill in *in situ* settings. *Sport Exerc. Perform. Psychol.* 4, 254–267. doi: 10.1037/spy 0000044

Müller, S., Brenton, J., Dempsey, A. R., Harbaugh, A. G. Y., and Reid, C. (2015b). Individual differences in highly skilled visual perceptual-motor striking skill. *Attention Percept. Psychophys.* 77, 1726–1736. doi: 10.3758/s13414-015-0876-7

Müller, S., Lalović, A., Dempsey, A. R., Rosalie, S., and Harbaugh, A. G. (2014). Pick-up of early visual information to guide kinetics and kinematics within a group of highly skilled baseball batters. *Percept. Motor Skills* 119, 347–362. doi: 10.2466/30.PMS.119c21z9

Muraskin, J., Sherwin, J., and Sajda, P. (2013). "A system for measuring the neural correlates of baseball pitch recognition and its potential use in scouting and player development," in *Proceedings of the MIT/Sloan Sports Analytics Conference.* (Boston, MA).

NCAA (2015). *Baseball Ranking Summary.* Available onine at: http://web1.ncaa.org/stats/StatsSrv/rankings

Panas, L. (2010). *Beyond Batting Average.* Available online at:https://www.lulu.com/

Park, J. L., Fairweather, M. M., and Donaldson, D. I. (2015). Making the case for mobile cognition: EEG and sports performance. *Neurosci. Biobehav. Rev.* 52, 117–130. doi: 10.1016/j.neubiorev.2015.02.014

Paull, G., and Glencross, D. (1997). Expert perception and decision making in baseball. *Int. J. Sport Psychol.* 28, 35–56.

Razer, A., Blair, L., and Fadde, P. (2015). "Accelerating expert noticing in classroom teaching, nursing, and academic coaching," in *Paper Presented at the Association for Educational Communications and Technology Conference.* (Indianapolis, IN).

Richey, R. C., Klein, J. D., and Tracey, M. W. (2011). *The Instructional Design Knowledge Base: Theory, Research, and Practice.* New York, NY: Routledge.

Tokmak, H. S., Baturay, H. M., and Fadde, P. (2013). Applying the context, input, process, product evaluation model for evaluation, research, and redesign of an online master's program. *Int. Rev. Res. Open Distributed Learn.* 14, 273–293.

Turner, B. (2015). *Jason Giambi Launches Interactive Hitting Simulator with Virtual Reality.* SportTechie. Available online at: http://www.sporttechie.com/2015/10/28/jason-giambi-launches-interactive-hitting-simulator-virtual-reality/

Ward, P., Ericsson, K. A., and Williams, A. M. (2013). Complex perceptual-cognitive expertise in a simulated task environment. *J. Cogn. Eng. Decis. Mak.* 7, l231–254. doi: 10.1177/1555343412461254

Ward, P., Farrow, D., Harris, K. R., Williams, A. M., Eccles, D. W., and Ericsson, K. A. (2008). Training perceptual-cognitive skills: can sport psychology research inform military decision training? *Mil. Psychol.* 20(Suppl. 1), S71–S102. doi: 10.1080/08995600701804814

Weinberg, N. (2014). *How to Evaluate a Hitter, Sabermetrically. Beyond the Box Score.* Available online at: http://www.beyondtheboxscore.com/2014/5/26/5743956/sabermetrics-stats-offense-learn-sabermetrics

White, P. (2014). *Hard Knock Life: Astros Turn to Plate Discipline for Success.* USA Today. Available online at: http://www.usatoday.com/story/sports/mlb/2014/06/03/houston-astros-build-plate-discipline-from-the-ground-up/9934613/

Williams, A. M., and Grant, A. (1999). Training perceptual skill in sport. *Int. J. Sport Psychol.*30, 194–220.

Williams, A. M., and Ward, P. (2003). "Perceptual expertise: development in sport," in *Expert Performance in Sports: Advances in Research on Sport Expertise,* eds J. L. Starkes and K. A. Ericsson (Champaign, IL: Human Kinetics), 219–247.

Williams, A. M., and Ericsson, K. A. (2005). Perceptual-cognitive expertise in sport: some considerations when applying the expert performance approach. *Hum. Mov. Sci.* 24, 283–307. doi: 10.1016/j.humov.2005.06.002

Williams, T., and Underwood, J. (1971). *The Science of Hitting.* New York, NY: Simon and Schuster.

Yin, R. K. (2014). *Case Study Research: Design and Methods.* 5th Edn. Thousand Oaks, CA: Sage.

## APPENDIX: BASEBALL PRIMER

A baseball game features nine players in each team's lineup. The teams alternate turns batting and playing the field for each of nine scheduled innings or until a winner is determined; there are no ties in baseball. A team can score runs (points) only when batting. As shown in **Figure A1**, the game is played on a field vaguely shaped like a diamond. Home plate and three bases define the infield, all 90 feet apart and separated by wide dirt running paths. When a team is in the field, the pitcher stands in the middle of the diamond (position 1 in **Figure A1**) and throws pitches to the catcher (position 2) who is positioned behind home plate, approximately 60 feet from the pitcher. Four infielders assume positions (3, 4, 5, and 6) loosely associated with each base. Three outfielders (positions 7, 8, and 9) patrol the expanse of grass between the infield and the outfield fence. The fence can be symmetrical or have odd shapes, such as legendary Fenway Park that was built to fit on a city block in Boston. The distance from home plate to the outfield fence varies and is not officially defined but is typically between 300 and 400 feet from home plate. When a pitched ball is batted into the fair playing area fielders can record an out by either catching the ball in the air or fielding the ball on the ground and throwing it to a base before the runner reaches the base. The defense must record three outs to end the inning.

When a team is on offense, the nine players become batters. Each batter has a turn to go to home plate and face the pitcher. Positioned behind the batter and catcher is an umpire who judges whether pitches are strikes or balls. The batter attempts to get a hit by batting a pitched ball so that it is not caught or he runs to base before a fielder's throw reaches the base. If he gets safely
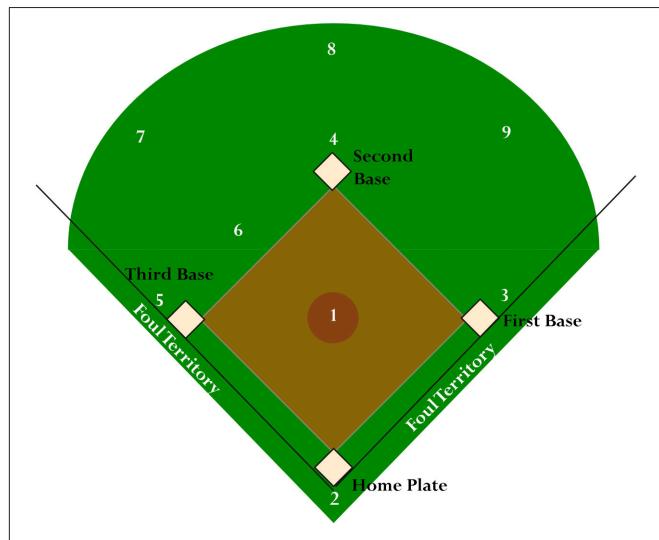


**FIGURE A1 | Diagram of baseball diamond (courtesy:cliparthut.com/).**

to first base, he has hit a single. Reaching second base safely represents a double, third base a triple, and rounding all of the bases is a home run (usually attained by hitting the ball over the outfield fence.) Every batter who reaches home plate scores a run. Batters have three strikes to put a pitch in play, being called out for swinging and missing on the third strike or "taking" (refraining from swinging at) a pitch that the plate umpire determines was in the strike zone. If the batter refrains from swinging at four pitches that the umpire determines are "balls" outside of the strike zone then the batter earns a base-on-balls, or walk, to first base.

# Macrocognition: From Theory to Toolbox

*Gary Klein[1]\* and Corinne Wright[2]*

[1] *MacroCognition LLC, Washington, DC, USA,* [2] *ShadowBox LLC, Dayton, OH, USA*

We trace several trajectories—the evolution of field-based decision making models in the mid-1980s to the formation of the Naturalistic Decision Making movement in 1989, then the further broadening of NDM into Macrocognition in 2003, and finally the transition from macrocognitive models into a set of methods and tools to boost cognitive performance.

Keywords: NDM, ShadowBox, cognitive, decision making, complexity

## STAGE 1: NATURALISTIC DECISION MAKING

During the 1980s, several researchers (Rasmussen, 1985; Cohen, 1986; Beach and Mitchell, 1987; Klein, 1989; Noble, 1989) independently began investigating the nature of decision making in work settings as opposed to laboratory, controlled settings. The NDM movement was catalyzed by a program established by Judith Orasanu at the Army Research Institute for the Behavioral and Social Sciences. Orasanu's program funded several of the investigators and also brought them together at periodic program reviews. A critical mass formed, leading to a 1989 workshop to prepare a book describing the NDM perspective. Approximately 30 researchers were invited to the meeting, including representatives from the US Army, Navy, and Air Force. The Navy was particularly interested in the topic because the Vincennes shoot down had occurred just a year earlier—an advanced AEGIS cruiser had shot down an Iranian commercial airliner, mistaking it for an attacking F-14. The Navy was shortly to initiate its own program of naturalistic decision research, Tactical Decision Making Under Stress (TADMUS).

The 1989 workshop resulted in an edited book, *Decision making in action: Models and methods* (Klein et al., 1993). The central NDM theme was to stud decision making under complex conditions, with vague goals, organizational constraints, high stakes, and levels of experience not easily captured in controlled laboratory settings (see **Figure 1**).

The term "Naturalistic Decision Making" was coined at the 1989 workshop, mirroring a related emerging topic of interest in the psychology of learning, *Naturalistic Memory*, initiated by Ulric Neisser. Naturalistic Memory encompassed topics of everyday memory, autobiographical memory, and practical memory (Gruneberg et al., 1978; Neisser, 1982).

Yet Naturalistic Memory quickly faded, even though Neisser was an iconic figure whose 1967 book *Cognitive Psychology* had helped to establish a new discipline (Neisser, 1967). Why did this happen? One possibility is that because Neisser was so famous, his suggestion of studying memory under non-controlled conditions was seen as blasphemous. In 1989 the American Psychologist devoted a special issue to allow critics to explain why Neisser's project was misguided (e.g., Banaji and Crowder, 1989—"The bankruptcy of everyday memory"). Laboratory researchers ridiculed the notion that anything of use could be learned from studying memory in natural settings. It is jarring to read their comments, aimed at one of the giants of the field, insisting that the new methodology should not even be explored.

In contrast, NDM already had generated an important discovery. Klein et al. (1986) and Klein (1989) described how people were able to make decisions under time pressure and uncertainty—the Recognition-Primed Decision (RPD) model. This finding parried any criticism that there was

FIGURE 1 | NDM variables (Klein et al., 1993). Illustrated by David Sweeney.



FIGURE 2 | Macrocognitive functions and processes.
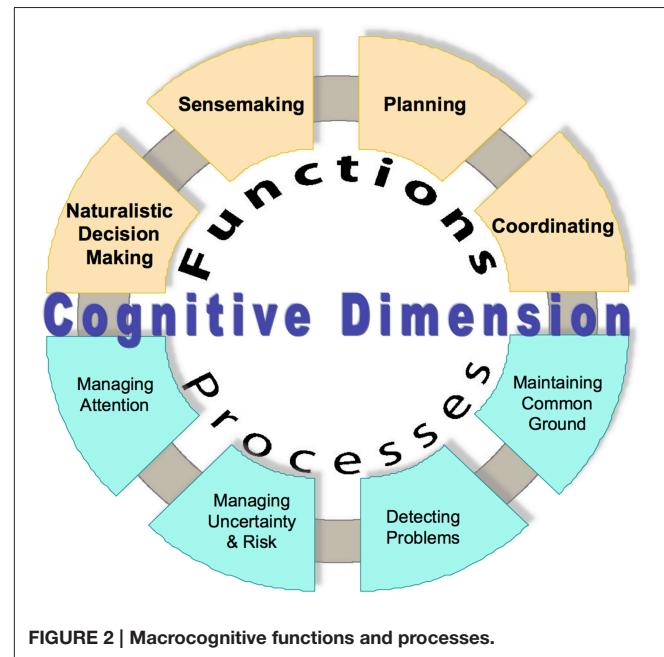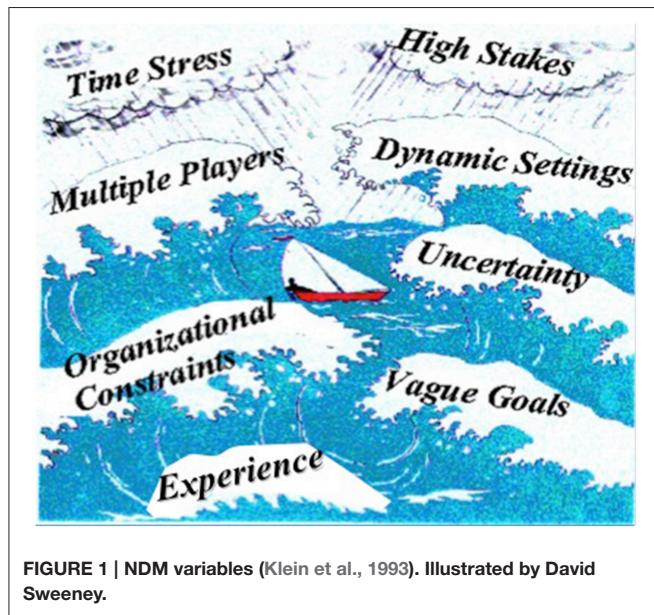
nothing to be learned from studying decision making in a natural setting. The RPD model accounted for 80–90% of the decisions that firefighters made. And the RPD model could never have been discovered under laboratory conditions because the RPD model depended on experience that took 10–20 years to develop. Laboratory-based decision research gave college sophomores unfamiliar tasks in order to avoid any contaminating effects of prior experience that might add unwanted variability to the results. Findings supporting the RPD model were replicated by different research teams in different contexts: military leaders (Schmitt and Klein, 1999), firefighters, (Keren et al., 2013), and managers of offshore oil drilling platforms (Skriver, 1998). The RPD model itself was tested and received empirical support (Klein et al., 1995; Johnson and Raab, 2003).

Naturalistic Decision Making researchers study how people actually make decisions. The word "actually," may discomfort laboratory researchers who can reasonably argue that even college students performing unfamiliar tasks are "actually" making decisions. However, many compromises have to be made to perform controlled experiments. The restriction on context, the absence of meaningful consequences, the use of tasks with well-defined goals, and particularly the elimination of expertise in studies presenting unfamiliar tasks, all raise doubts about whether the findings of these studies can be generalized to natural settings. Laboratory researchers can counter that the lack of controlled conditions also raise doubts about the results of NDM studies. We are not arguing that either tradition is the correct one. We merely assert that NDM projects offer unique opportunities for discoveries.

There was some criticism of NDM from laboratory-based decision researchers. Lipshitz et al. (2001) published a lead article in the *Journal of Behavioral Decision Making*, and an unprecedented number of researchers wrote commentaries, 16 in all, centered around the theme that NDM needed to mature as a discipline and establish more rigorous and controlled methods of

investigation. The NDM community viewed these criticisms as misguided. If NDM researchers followed the critics' suggestions and performed studies under controlled conditions, they would no longer be doing NDM work. The scientific method begins with observing the phenomenon of interest, which is the core of NDM research.

Unlike Naturalistic Memory, which quickly faded, NDM thrived after the first NDM workshop in 1989. Thus, far, a total of 12 NDM conferences have been held. A Cognitive Engineering and Decision Making technical group, drawing on the NDM community, was established in 1995 within the Human Factors and Ergonomics Society. This group now has its own *Journal of Cognitive Engineering and Decision Making*.

## STAGE 2: MACROCOGNITIVE MODELS

Because of the success of naturalistic inquiry into decision making, NDM researchers quickly began applying this approach to other cognitive phenomena, such as planning (e.g., Klein, 2007a,b), sensemaking (e.g., Klein et al., 2006), and uncertainty management (e.g., Lipshitz and Strauss, 1997), as well as the development of expertise itself (e.g., Klein, 1997). **Figure 2** illustrates the range of cognitive functions and processes addressed by macrocognitive models (Klein et al., 2003).

Klein et al. (2000) differentiated macrocognition and microcognition. They defined macrocognition as the study of cognitive processes affecting people such as firefighters, pilots, nurses, and others who had to wrestle with difficult dilemmas in complex settings under time pressure and uncertainty. Microcognition was the study of the components of thinking such as working memory, and serial vs. parallel attentional processes. Other researchers had used the term macrocognition (e.g., Cacciabue and Hollnagel, 1995) in papers, but had not

identified it as a separate field of study, an expansion of the NDM enterprise (for a fuller history, see Hoffman and McNeese, 2009).

The NDM community has now expanded its perspective beyond decision making, to cover the variety of macrocognitive models and to perform naturalistic studies of cognitive processes and variables. Sometimes, macrocognitive studies are performed under controlled conditions and the studies often involve the control and manipulation of variables. But the core of the work examines cognitive processes in complex contexts—in the context of the work environment. Macrocognitive studies usually address expertise, how it develops, what constitutes it, how it is used to perform challenging tasks. Sometimes researchers will investigate novices, to see how they differ from experts, how they approach tasks, and where they struggle.

The criticisms of NDM that have been raised by laboratory scientists also will apply to macrocognition. It is less concerned with testing hypotheses than with formulating useful models and theories. It is less concerned with precision than with plausibility. It is less concerned with normative or "rational" models than with descriptive models. It wallows in messy variables such as wicked problems with ill-defined goals, team and organizational constraints, uncertainty, and high stakes. It studies tasks for which there are no correct solutions, making it difficult to evaluate performance. Guilty as charged. These are the conditions in which we live and work, and macrocognitive research attempts to better understand them. **Figure 1**, a diagram originally formulated for NDM equally well illustrates the variables of interest for macrocognition.

Klein (2015) described the impact of the NDM/macrocognitive perspective, by cataloging the way this perspective has changed so many core beliefs previously held in the basic and applied communities.

We no longer claim that the only way to make a good decision is to generate several options and compare them to pick the best one (experienced decision-makers can draw on patterns to handle time pressure and never even compare options; Klein, 1989; Hoffman, 1992). We no longer believe that expertise is based on learning rules and procedures (it primarily depends on tacit knowledge, Klein and Hoffman, 1993). We no longer believe that projects must start with a clear description of the goal (many projects involve wicked problems and ill- defined goals, Hoffman, 2007). We no longer believe that people make sense of events by building up from data to information to knowledge to understanding (experienced personnel use their mental models to define what counts as data in the first place, Skriver et al., 2004; Schraagen et al., 2008). We no longer believe that insights arise by overcoming mental sets (they also arise by detecting contradictions and anomalies and by noticing connections, Klein, 2013). We no longer believe that we can reduce uncertainty by gathering more information (performance seems to go down when too much information is gathered—Uncertainty can stem from inadequate framing of data, not just from the absence of data, Cannon-Bowers et al., 1993; Omodei et al., 2005; Flin et al., 2008; Grossman et al., 2014). We no longer believe that we can improve performance by teaching critical thinking precepts such as listing assumptions (too often the flawed assumptions are ones we are not even aware of and would never list, Klein, 2011; Stanton et al., 2011; Hoffman et al., 2014).

Whereas the behavioral decision making community focuses on human limitations and seeks ways to reduce biases and mistakes, the NDM community, as it performs macrocognitive research, focuses on human capabilities and regards good performance as much more than the absence of mistakes. Good performance is also about discoveries and insights; it is about the strengths of decision makers, and the importance of experience. Experience serves a variety of functions including a larger repertoire of patterns and associated actions, a richer mental model of how things work to support inferential reasoning and sensemaking for diagnosis and anticipation.

## STAGE 3: MACROCOGNITIVE METHODS AND TOOLS

Macrocognitive models lend themselves to a set of methods and tools that can be used in cognitively challenging activities. This third stage is about compiling a toolbox, not to do research, but to enhance performance.

A recent study illustrates how a macrocognitive perspective provides a unique diagnosis of a problem, not by blaming those committing the error (those on the sharp end, as James (Reason, 1990), would put it), but by investigating how conscientious employees could be making poor decisions (Multer et al., 2015; Safar et al., 2015). The problem was railroad crashes caused by locomotive engineers who failed to stop despite clear signals warning them of danger. It seemed obvious that the locomotive engineers were getting distracted or were failing to pay adequate attention. Strategies were devised to help the engineers, including a "Keep the Focus" program. However, Multer et al. (2015) viewed the inattention/distraction issue as a *symptom* of the problem, not as the *source* of the problem. They investigated the reasons for the inattentiveness, using interviews and field observations. They examined the sociotechnical context of the "Signals Passed at Danger" phenomenon. One finding was that each signal had several lights, including a red light, but the red light was always on! Thus, the red light provided no information. The other lights signaled whether to stop or proceed with caution.

Another finding was that central stations had outgrown their original design because railway traffic had increased. Space became more cramped, the switching arrangements had become more complex, and the viewing angles became more ambiguous as different signals were moved closer and closer, to the point that the engineers were not always sure which signal pertained to which line. Worse, because trains were now longer, signals were sometimes placed behind the train and out of sight of the engineers in the locomotive cab.

Holtrop et al. (2015) used Cognitive Task Analysis in the domain of healthcare. They were sending healthcare practitioners—nurses, aides, etc.—into the community to work with patients who had chronic illnesses such as cardiac disease and diabetes. But the effort was running into barriers because each clinic and practice had its own decision making style. So the project added a Cognitive Task Analysis training piece, a two-day workshop to train the outreach personnel in CTA methods in order to overcome the differences in decision making. The

training was highly successful, and the recommendations gained greater acceptance. Accordingly, the healthcare practitioners became advocates for front-end CTA analysis prior to initiating any new effort.

Many domains depend on training to get employees up to speed, but the training usually centers on rules and procedures. What is missing is a concern for the tough decisions employees will have to make once they complete their training: the difficult sensemaking they will face when confronted with ambiguous cues and erroneous data, the challenging problem detection when things are just starting to go wrong. The field of macrocognition is well suited for addressing cognitive training requirements.

Hoffman et al. (2014) provided an important resource—a compilation of best practices for accelerating the development of expertise. They identified strategies for practice and feedback, transfer, and retention, and also addressed team training issues. Expertise is central to all the macrocognitive processes. Tactics for speeding up expertise are essential macrocognitive tools.

The design of new systems, and the modification of existing systems, can benefit greatly from a macrocognitive perspective, and a variety of methods have emerged for injecting cognition into the design process (e.g., Militello and Hutton, 1998; Militello et al., 2010; Militello and Klein, 2013) and for making automation a team player (Klein et al., 2004). The intent here is to develop macrocognitive work systems.

One last example of a macrocognitive tool is the ShadowBox approach. ShadowBox® training is a scenario-based way to enable trainees to see the world through the eyes of experts without the experts having to be present. One of the bottlenecks of expert feedback is that there is a limited number of experts and their time is jealously guarded. ShadowBox presents challenging scenarios and intersperses decision points at which the participant is asked to rank a given set of options. These may be options about which course of action to choose, which goal to prioritize, which cues to monitor carefully, or which pieces of information to gather. A participant ranks the options and writes his/her rationale for the rankings. In preparation for the training session, a panel of subject-matter experts also has rank ordered the options and provided their rationale statements. At this point, the experts are no longer needed. The expert rankings are combined, and the rationale statements are synthesized. Once the participant provides rankings and rationale, he/she sees what the panel of experts ranked, and sees the experts' rationale, noticing what he/she had missed. In this way, the participant gets to see the scenario through the eyes of the experts, and gets a sense of the experts' mental models.

Hintze (2008), a Battalion Chief with the New York Fire Department originated the ShadowBox concept. Klein et al.

(2013) elaborated on the ShadowBox concept. While it is primarily a means of providing cognitive training, it also can be used as a knowledge management tool to capture the wisdom of experts in the form of the rankings and reasons they provide. A third use of ShadowBox is for assessment, using a participant's rankings and rationale to evaluate competence. And a fourth use is for better teamwork. Here, team members identify how they would react at critical moments in a scenario, and they predict how their partners would react. Then each partner sees what the other would do, and what the other expected. In this way, teams can get better at predicting what the others will do; predictability is essential to team coordination. A fifth use of ShadowBox is to support leadership. For a given scenario, the leader or supervisor acts in the place of the panel of experts, and gets to see how the subordinates would act, and what their reasoning was; the idea is to strengthen the calibration of the leader and the subordinates. The subordinates don't have to agree with the supervisor, but they do have to understand what the supervisor expects and how the supervisor interprets the situation.

## CONCLUSION

The NDM framework has developed over the past 30 years, shaping the thinking and capabilities of a community of researchers and practitioners. The NDM community is now engaged in studying macrocognitive phenomena and developing methods for supporting these functions and processes. This work has shaped our thinking about cognitive processes such as decision making, sensemaking, and problem detection that are engaged in complex and uncertain environments. It has shaped our capabilities for training, decision support systems, and system design. NDM researchers typically work with domain specialists performing complex and challenging tasks. Accordingly, the methods and the models are especially suited to applications and are grounded in the variables that matter the most in "natural" conditions.

## AUTHOR CONTRIBUTIONS

GK is the lead author and took primary responsibility in preparing this manuscript. CW is the second author and helped with the writing and preparation.

## ACKNOWLEDGMENTS

## REFERENCES

Banaji, M. R., and Crowder, R. G. (1989). The bankruptcy of everyday memory. *Am. Psychol.* 44, 1185–1193. doi: 10.1037/0003-066X.44.9.1185

Beach, L. R., and Mitchell, T. R. (1987). Image theory: principles, goals, and plans in decision making. *Acta Psychol.* 66, 201–220. doi: 10.1016/0001-6918(87)90034-5

Cacciabue, P. C., and Hollnagel, E. (1995). *Simulation of Cognition: Applications.* Hillsdale, NJ: Lawrence Erlbaum Associates, 55–73.

Cannon-Bowers, J. A., Salas, E., and Converse, S. A. (1993). "Shared mental models in expert decision making," in *Individual and Group Decision Making: Current Issues*, ed N. J. Castellan (Hillsdale, NJ: Lawrence Erlbaum Associates), 221.

Cohen, M. S. (1986). "An expert system framework for non-monotonic reasoning about probabilistic assumptions," in *Uncertainty in Artificial Intelligence:*

*Machine Intelligence and Pattern Recognition*, Vol. 4. eds L. N. Kanal and J. F. Lemmer (North Holland: Elsevier), 279–293.

Flin, R., O'Connor, P., and Crichton, M. (2008). *Safety at the Sharp End. A Guide to Non-Technical Skills*. Farnham, UK: Ashgate.

Grossman, R., Spencer, J. M., and Salas, E. (2014). "Enhancing naturalistic decision making and accelerating expertise in the workplace: training strategies that work," in *Judgment and Decision Making at Work*, eds S. Highhouse, R. S. Dalal, and E. Salas (New York, NY: Routledge), 277–325.

Gruneberg, M. M., Morris, P. E., and Sykes, R. N. (1978). *Practical Aspects of Memory*. London: Academic Press.

Hintze, N. R. (2008). *First Responder Problem Solving and Decision Making in Today's Asymmetrical Environment*. Doctoral dissertation, Monterey, California, Naval Postgraduate School.

Hoffman, R. R. (ed.). (1992). *The Psychology of Expertise: Cognitive Research and Empirical AI*. New York, NY: Springer-Verlag.

Hoffman, R. R. (ed.). (2007). *Expertise Out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*. Mahwah, NJ: Lawrence Erlbaum.

Hoffman, R. R., and McNeese, M. (2009). A history for macrocognition. *J. Cogn. Eng. Decis. Making* 3, 97–110. doi: 10.1518/155534309X441835

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., and Andrews, D. H. (2014). *Accelerated Expertise: Training for High Proficiency in a Complex World*. New York, NY: Taylor and Francis; Psychology Press.

Holtrop, J. S., Potworowski, G., Fitzpatrick, L., Kowalk, A., and Green, L. A. (2015). Understanding effective care management implementation in primary care: a macrocognition perspective analysis. *Implement. Sci.* 10, 122. doi: 10.1186/s13012-015-0316-z

Johnson, J. G., and Raab, M. (2003). Take the first: option-generation and resulting choices. *Organ. Behav. Hum. Decis. Process.* 91, 215–229. doi: 10.1016/S0749-5978(03)00027-X

Keren, N., Bayouth, S. T., Franke, W. D., and Godby, K. M. (2013). "Examining the effect of level of stress on firefighters' time-to-decision in virtual reality," in *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (San Diego, CA), 299–303.

Klein, D. E., Klein, H. A., and Klein, G. (2000). "Macrocognition: linking cognitive psychology and cognitive ergonomics," in *Proceedings of 5th International Conference Human Interactions with Complex Systems* (Urbana-Champaign: University of illinois at Urbana-Champaign, The Beckman Institute; U.S. Army Research Laboratory, Advanced Displays & Interactive Displays; Federated Laboratory Consortium), 173–177.

Klein, G. (1997). Developing expertise in decision making. *Think. Reason.* 3, 337–352. doi: 10.1080/135467897394329

Klein, G. (2007a). Flexecution as a paradigm for replanning, Part 1. *IEEE Intell. Syst.* 22, 79–83. doi: 10.1109/MIS.2007.4338498

Klein, G. (2007b). Flexecution as a paradigm for replanning, Part 2: understanding and supporting flexible execution. *IEEE Intell. Syst.* 22, 108–112. doi: 10.1109/MIS.2007.107

Klein, G. (2013). *Seeing What others Don't: The Remarkable Ways We Gain Insights*. New York, NY: Public Affairs.

Klein, G. (2015). Reflections on applications of naturalistic decision making. *J. Occup. Organ. Psychol.* 88, 382–386. doi: 10.1111/joop.12122

Klein, G., Hintze, N., and Saab, D. (2013). "Thinking inside the box: the ShadowBox method for cognitive skill development," in *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013)* (Paris: Arpege Science Publishing).

Klein, G., Moon, B., and Hoffman, R. R. (2006). Making sense of sensemaking 2: a macrocognitive model. *IEEE Intell. Syst.* 21, 88–92. doi: 10.1109/MIS.2006.100

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., and Hollnagel, E. (2003). Macrocognition. *IEEE Intell. Syst.* 18, 81–85. doi: 10.1109/MIS.2003.1200735

Klein, G., Wolf, S., Militello, L., and Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organ. Behav. Hum. Decis. Process.* 62, 63–69.

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

Klein, G. A. (1989). "Recognition-primed decisions," in *Advances in Man-Machine Systems Research*, Vol. 5. ed W. B. Rouse (Greenwich, CT: JAI Press), 47–92.

Klein, G. A. (2011). Critical thoughts about critical thinking. *Theor. Issues Ergonom. Sci.* 12, 210–224. doi: 10.1080/1464536X.2011.564485

Klein, G. A., Calderwood, R., and Clinton-Cirocco, A. (1986). "Rapid decision making on the fire ground," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 30* (Dayton, OH: SAGE Publications), 576–580.

Klein, G. A., and Hoffman, R. (1993). "Seeing the invisible: perceptual/cognitive aspects of expertise," in *Cognitive Science Foundations of Instruction*, ed M. Rabinowitz (Mahwah, NJ: Lawrence Erlbaum), 203–226.

Klein, G. A., Orasanu, J., Calderwood, R., and Zsambok, C. (1993). *Decision Making in Action. Models and Methods*. Norwood, NJ: Ablex.

Lipshitz, R., Klein, G., Orasanu, J., and Salas, E. (2001). Focus article: taking stock of naturalistic decision making. *J. Behav. Decis. Mak.* 14, 331–352. doi: 10.1002/bdm.381

Lipshitz, R., and Strauss, O. (1997). Coping with uncertainty: a naturalistic decision-making analysis. *Org. Behav. Hum. Decis. Process.* 69, 149–163. doi: 10.1006/obhd.1997.2679

Militello, L. G., Dominguez, C. O., Lintern, G., and Klein, G. (2010). The role of cognitive systems engineering in the systems engineering design process. *Syst. Eng.* 13, 261–273. doi: 10.1002/sys.20147

Militello, L. G., and Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics* 41, 1618–1641. doi: 10.1080/001401398186108

Militello, L. G., and Klein, G. (2013). "Decision-centered design," in *Oxford Handbook of Cognitive Engineering*, eds J. Lee and A. Kirlik (Oxford, UK: Oxford University Press), 261–271.

Multer, J., Safar, H., and Roth, E. (2015). "Understanding how signals passed at danger occur through the lens of sociotechnical systems," in *Paper Presented at the 5th International Rail Human Factors Meeting* (London, UK).

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Neisser, U. (1982). *Memory Observed*. New York, NY: W.H. Freeman.

Noble, D. (1989). *Application of a Theory of Cognition to Situation Assessment*. Vienna, VA: Engineering Research Associates.

Omodei, M. M., McLennan, J., Elliott, G. C., Wearing, A. J., and Clancy, J. M. (2005). "More is better? A bias toward overuse of resources in naturalistic decision-making settings," in *How Professionals Make Decisions*, eds H. Montgomery, R. Lipshitz, and B. Brehmer (Mahwah, NJ: Lawrence Erlbaum), 29–40.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decisionmaking and system management. *Syst. Man Cybern. IEEE Trans.* 2, 234–243.

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

Safar, H., Multer, J., and Roth, E. (2015). *An Investigation of Passing Stop Signals at a Passenger Railroad*. Technical Report, U.S. Department of Transportation, Federal Railroad Administration. No. DOT-VNTSC-FRA-15-A02.

Schmitt, J. F., and Klein, G. (1999). How we plan. *Mar. Corps Gazette* 83, 18–26.

Schraagen, J. M. C., Militello, L., Ormerod, T., and Lipshitz, R. (eds.). (2008). *Naturalistic Decision Making and Macrocognition*. Aldershot, UK: Ashgate.

Skriver, J. (1998). *Emergency Decision Making on Offshore Installations*. Doctoral dissertation, University of Aberdeen.

Skriver, J., Martin, L., and Flin, R. (2004). "Emergency decision making," in *Psychological Investigations of Competence in Decision Making*, eds K. Smith, J. Shanteau, and P. Johnson (Cambridge, UK; New York, NY: Cambridge University Press), 124–148.

Stanton, N. A., Wong, W., Gore, J., Sevdalis, N., and Strub, M. (2011). Critical thinking. *Theor. Issues Ergonom. Sci.* 12, 204–209. doi: 10.1080/1464536X.2011.564479

# Advantages of publishing in Frontiers

**OPEN ACCESS**

Articles are free to read, for greatest visibility

**COLLABORATIVE PEER-REVIEW**

Designed to be rigorous – yet also collaborative, fair and constructive

**FAST PUBLICATION**

Average 85 days from submission to publication (across all journals)

**COPYRIGHT TO AUTHORS**

No limit to article distribution and re-use

**TRANSPARENT**

Editors and reviewers acknowledged by name on published articles

**SUPPORT**

By our Swiss-based editorial team

**IMPACT METRICS**

Advanced metrics track your article's impact

**GLOBAL SPREAD**

5'100'000+ monthly article views and downloads

**LOOP RESEARCH NETWORK**

Our network increases readership for your article

**Find us on**