Teamwork in Extreme Environments

Olivia Brown
BSc, MSc

This thesis is submitted in partial fulfilment for the degree of Doctor of Philosophy

May 2020

Psychology Lancaster University

Postgraduate Research Degrees
FORM: PERMISSION TO
SUBMIT A THESIS IN ALTERNATIVE FORMAT

Name of Candidate: (PLEASE PRINT) Olivia Brown

Date of Registration: 01/10/2016

Registration Number:

Division and Faculty: PSYCHOLOGY - FST

Research Degree Title: Teamwork in extreme environments

Date Thesis due for Submission: End of March/Start of April

This is to confirm that the above named student has been granted permission to submit a PhD or MPhil thesis in Alternative Format.

Name(s) of Supervisor(s) (PLEASE PRINT): STACEY CONCHIE

Signature(s): S Conchie

Signature of HoD, or Director of Graduate Studies: [signature]

DATE of FORM SUBMISSION: 5/12/19

NB: A copy of this form must be held in the Psychology Department's Postgraduate Office, and a copy must be submitted with the thesis.

Declaration

I hereby declare that the work in this thesis has not been submitted in whole or in part, for the award of a higher degree at this or any other University. Further, this thesis is a product of my own work and the content of this thesis reflects my own thinking. The research in this thesis was completed under the supervision of Dr Nicola Power, Professor Emma Barrett and Dr Stacey Conchie.

Olivia Brown, B.Sc., M.Sc.

Lancaster University

Statement of Authorship

Chapter II - Brown, O., Power, N. & Conchie, S. (2020). Immersive Simulations with Extreme Teams. *Organizational Psychology Review. **

Conception of review paper: Olivia Brown, Nicola Power, Stacey Conchie
Drafting the manuscript: Olivia Brown
Revising the manuscript: Olivia Brown, Nicola Power, Stacey Conchie
Contribution of principal author: 85%

Chapter III - Brown, O., Power, N. & Conchie, S. (2020). Communication and coordination across time in the multi-team system response to a simulated terrorist incident. Submitted to *Journal of Occupational and Organizational Psychology.*

Conception and design of study: Olivia Brown, Nicola Power
Acquisition of data: Olivia Brown, Nicola Power
Data analysis: Olivia Brown
Data interpretation: Olivia Brown, Nicola Power
Drafting the manuscript: Olivia Brown
Revising the manuscript: Olivia Brown, Nicola Power, Stacey Conchie
Contribution of principal author: 85%

Chapter IV - Brown, O., Power, N. & Conchie, S. (2020). Multi-team systems in crises; does Familiarity matter? Submitted to *Small Group Research.*

Conception and design of study: Olivia Brown, Nicola Power
Acquisition of data: Olivia Brown
Data analysis: Olivia Brown
Data interpretation: Olivia Brown, Stacey Conchie
Drafting the manuscript: Olivia Brown
Revising the manuscript: Olivia Brown, Nicola Power, Stacey Conchie
Contribution of principal author: 85%

Chapter V - Brown, O., Barrett, E. & Power, N. (2020). Monitoring team cohesion over time in expedition teams; the role of daily events and team composition. In *Proceedings of the 14th International Naturalistic Decision-Making Conference, San Francisco, U.S.A.**

Conception and design of study: Olivia Brown, Emma Barrett
Acquisition of data: Olivia Brown, Emma Barrett
Data analysis: Olivia Brown

Data interpretation: Olivia Brown, Emma Barrett
Drafting the manuscript: Olivia Brown
Revising the manuscript: Olivia Brown, Emma Barrett, Nicola Power
Contribution of principal author: 85%

*published at the time of thesis submission

Signed confirmation that the above information relating to the contributions of Authors is correct:

Olivia Brown

Dr Nicola Power

Professor Emma Barrett

Dr Stacey Conchie

Acknowledgements

Firstly, I must thank my amazing supervisory team for all of their advice and guidance throughout this PhD. Nikki, thank you for guiding me throughout this process, expanding my thinking and for all the opportunities you have afforded me. Emma, thank you for always challenging my thinking, helping me to build confidence in my ideas and for always making me feel seen, no matter what crowd we were in. Stacey - I can't thank you enough for your support and guidance these last few months, you've gone above and beyond and been instrumental in getting this thesis over the line. Also, to Nathan, your research on extreme environments is inspiring and the conversations we've had over the years have really contributed to my thinking in writing this thesis.

A special thank you to CREST for funding this research and providing me with invaluable opportunities to share my research with relevant stakeholders along the way. It's been wonderful to be surrounded by a community of like-minded researchers. To the wonderful CREST team up in Lancaster; Nikki, Becky and Jen - thank you for everything you've done to make this whole experience as enjoyable as possible and for always stepping in to save the day when things didn't go to plan!

To my friends who have been there during this whole process, thank you. Frankie – we started this PhD journey together and it's been an absolute lifesaver having you there to talk things through and share ideas. And to Mo, thanks for the kindness and encouragement, especially during the final few months. And of course, a special thanks to Root coffee cafe for keeping me well caffeinated throughout.

To my mum and dad – where do I start. This PhD would never have come to being if it wasn't for your relentless support. I still remember A Level results day and the crippling disappointment of missing my grades. What a blessing to have you both by my side to pick me up and push me on. I am forever grateful for having you as my parents - thank you. And to my sister Rosie, thank you for believing in me, always making me laugh and never letting me get too far ahead of myself.

Finally, thank you to my darling Holly. You have been my absolute rock. I owe so much of this thesis to your endless support and love. I count myself lucky every day to have you in my corner and I'm so excited for everything we have planned this year and for many years to come.

Contents

List of Figures

List of Tables

List of Appendices

Abstract

Teams are relied upon in extreme and challenging environments in which there are considerable demands and failures can have severe consequences. Despite an increased interest in extreme teams, empirical research remains limited. Moreover, whilst the literature differentiates between extreme and non-extreme teams, it rarely distinguishes between different *types* of extreme teams. In this thesis, I argue extreme teams can be differentiated into multi-team systems (MTS) and teams in isolated, confined environments (ICE). I draw on contextual challenges present in different types of extreme environments to examine what factors support teamwork in emergency response teams (MTS) and expedition teams (ICE). In doing so, I identify methodological and analytical approaches suitable for researching extreme teams (*Chapter II* and *V)*.

MTS often form quickly in the response to emergencies. This creates challenges in establishing communication channels and managing conflicting objectives across inter-agency partners who may have limited experience working with one another. To address these challenges, I explored how teams communicated and coordinated in crises and if this is influenced by team member familiarity. Data were collected from immersive simulations with commanders in the emergency services and students. Mixed methods analysis showed how team processes changed across time (*Chapter III)* and how familiarity (*Chapter IV)* alleviated some of the challenges of working in MTS. In contrast to MTS, teams in ICE co-exist for pro-longed periods in hostile and remote settings. This creates challenges in maintaining team cohesion and balancing the personalities and characteristics of isolated individuals for pro-longed periods. Here, I used a diary methodology to track changes in cohesion over time and explore if fluctuations in cohesion are predicted by day-to-day events and the personality composition of teams (*Chapter V)*. Theoretical implications for the importance of

context in shaping team behaviours and practical implications for teams operating in extreme environments are provided.

Chapter 1

Introduction

Introduction

Extreme teams (ETs) operate in challenging environments which pose extraordinary physical, psychological and interpersonal demands on individuals (Manzey & Lorenz, 1998). Collectively these environments are referred to as extreme environments: atypical contexts (e.g., time pressure, level of uncertainty and demand) in which ineffective performance can have severe, potentially life or death consequences (Bell, Fisher, Brown & Mann, 2018). Examples of ETs include crews of long-duration space flights (Salas, Tannenbaum, Kozlowski, Miller, Mathieu & Vessey, 2015), teams in nuclear plant control rooms (Stachowski, Kaplan & Waller, 2009), medical emergency teams (Klein, Ziegert, Knight & Xiao, 2006), submarine command and control teams (Stanton & Roberts, 2018), mountaineering teams (Wickens, Keller & Shaw, 2015), Arctic teams (Kjaergaard, Leon & Fink, 2015) and emergency response teams (Power, 2018). Teamwork is especially important in extreme environments to promote safety and effective working (Hughes et al., 2016). This is emphasised by the fact that errors in extremes are attributed to failures in collaboration across members as opposed to individual deficiencies or endogenous challenges (e.g., urgency of situation) in the environment (Alison, Power, van den Heuvel & Waring, 2015a; Dwyer & Smith, 1991; Risser, Rise, Salisbury, Simon and Berns, 1999). Advancing understanding of how to support effective teamwork in ETs is therefore of paramount importance. Whilst there have been substantial advancements in the science of teamwork in the past two decades, much of the empirical research has been grounded in the study of conventional teams (e.g., business teams) (Mathieu, Hollenbeck, Van Knippenberg & Ilgen, 2017; Mathieu, Wolfson & Park, 2018). Accordingly, questions remain as to how to support effective teamwork in extreme and challenging environments.

A team is defined as two or more interrelated individuals, each with pre-defined roles, tasked with completing a common goal (Salas, Dickinson, Converse & Tannenbaum, 1992). In the past two decades there has been a resurgence in small group research, leading to a shift in focus from understanding interpersonal groups to understanding how collaborative teamworking can improve performance, adaptability and safety (Kozlowski & Chao, 2018; Maynard, Kennedy & Sommer, 2015; Salas, Cooke & Rosen, 2008). This has been driven, in part, by an increased emphasis on the role of the group within the workplace (Mathieu et al., 2018; Cross, Rebele & Grant, 2016). This research has been fruitful in facilitating the development of models and enhancing our understanding of teamwork. However, it has meant the majority of empirical research has been conducted with conventional teams, with a focus on improving work-place productivity (*see* Mathieu, Mayard, Rapp & Gilson, 2008; Mathieu et al., 2018; Wageman, Hackman & Lehman, 2005). This is problematic when we consider that the context in which ETs operate is far removed from that of conventional teams, presenting unique theoretical and methodological challenges (Bell et al., 2018; Maynard, Kennedy & Resick, 2018). For example, in comparison to conventional teams, ETs have complex structures and operate in high stakes, rapidly changing conditions that can impose a high level of stress on team members (Kozlowski & Chao, 2018; Schmutz, Lei, Eppich & Manser, 2018). In light of this, researchers have called for more empirical research on ETs to establish how known factors of teamwork operate in extreme environments (Vessey & Landon, 2017).

This call informed the two main goals of this thesis: (i) to explore what factors support effective teamwork in extremes, and (ii) to establish the methodological and analytical approaches suited to studying ETs. Given the severity of the consequences of teamwork errors in extremes (e.g., loss of life), identifying factors that support teamwork has important practical implications (Kozlowski & Chao, 2018). Findings may illustrate differences from what we have come to understand in conventional teams, suggesting the need to develop specified interventions to support teamwork in ETs. For example, whilst we know that familiarity promotes teamwork in conventional teams, there is little to no evidence as to how it might support teamwork in larger multi-agency teams during disaster response (Huckman, Staats & Upton, 2009; Turner, Thurlow, Baker, Northcutt, & Newman, 2019).

A further important contribution of this thesis was to examine the differences between different types of ETs (intra-extreme differences). Whilst the literature on ETs has tended to treat them as holistically different from conventional work teams, it rarely distinguishes between the *types* of extreme environments. However, this thesis argues that ETs can be differentiated into rapidly forming multi-team systems (MTS) (e.g., medical emergency teams, emergency response teams) and teams in isolated confined environments (ICE), who work together for a period of weeks or months (e.g., expeditions teams, teams in long distance spaceflight). In this thesis, two teams are studied, one from each sub-set of ETs: emergency response teams (an example of MTS) and expedition teams (an example of teams in ICE).

The second goal of this thesis was to establish the methodological and analytical approaches suited to studying different types of ETs. This involved a careful consideration of the methodologies best suited to study these complex, often hard-to-reach teams, posing unique challenges in comparison to conventional teams, which are easier to access and more conducive to studies with large sample sizes (Bell et al., 2018). Methods with ETs must be robust enough to facilitate rigorous research that can improve safety and future working, as well as contributing to wider theoretical understanding (Kozlowski, 2015). Further the methods and analytical techniques must be appropriate to the team studied. For instance, whilst simulation studies are suitable for researching emergency response teams, they are unlikely to be appropriate for expedition teams, who operate for long periods in inhospitable climates.

Overall, this thesis aims to offer insight into how teams operate in extremes. Whilst there exists many review articles and commentaries on ETs, there remains a paucity of empirical research *(see* Power, 2018; Golden, Chang & Kozlowski, 2019; Maynard et al., 2018; Vessey & Landon, 2017). To address this, I explore what factors support effective teamwork in different types of extreme environment. This thesis contributes to the wider theoretical understanding of how context shapes team behaviours by exploring why variables may be more important in certain types of ET and by examining teamwork over time. It also identifies potential avenues where practical solutions to improving teamwork in ETs may be developed.

The remainder of this Chapter is structured as follows. First, I outline the research context. The focus here is to (i) outline how ETs differ from conventional teams, and (ii) highlight differences between types of ET and why it is important to recognise these differences. From this, I identify the two key research questions that will be addressed in this thesis. To help ground this thesis in existing teamwork research, the Chapter proceeds to (i) outline theoretical models of teams and, (ii) provide an overview of the literature around two specific types of ETs: emergency response teams (MTS) and expedition teams (ICE). Doing so illustrates how teams are conceptualised in this thesis and identifies the factors that are examined empirically in subsequent Chapters.

Research Context

Research has made great strides in understanding how individuals are able to work together effectively (Mathieu et al., 2018). With this development has come several definitions of teams, with many researchers showing a preference to focus on context (e.g., in organisational contexts, [Kozlwoski & Bell, 2003]; in sports teams, [McEwan & Beauchamp, 2014]; in the military, [Orasanu & Salas, 1993]). Focusing on context can limit the extent to which these definitions can be operationalised. For example, definitions of organisational teams have included reference to the organisational context (e.g., leadership structures) which can limit and define the way in which team members interact (Kozlowski & Bell, 2003). Similarly, those focusing on military teams reference task conditions in which there is a high level of uncertainty and demand (Orasanu & Salas, 1993). As the research in this thesis involves two types of team that are not bound to one specific context (MTS and ICE), I have elected to adopt a broader definition of teams as: two or more interrelated individuals, each with specified roles, interacting dynamically, interdependently and adaptively to purse a common goal (Salas et al., 1992). This definition captures the common elements found in teams across contexts and distinguishes teams from small groups through specified roles and interdependent actions (Dyer, 1984; Kozlowski & Ilgen, 2006). Further, as this thesis concerns teams in extreme environments, I define extreme teams as: (i) teams that complete their tasks in environments with one or more contextual features that are extreme in level/kind (e.g., time pressure, stress) and (ii) in which ineffective performance has severe consequences (e.g., harm to team or others) (Bell et al., 2018).

For a team to function effectively, they must succeed in taskwork and teamwork (Burke, Wilson & Salas, 2003; Crawford & LePine, 2013; Salas, Cooke & Rosen, 2008). Taskwork refers to the completion of specific, work-related activities required to achieve team goals (Dinh & Salas, 2017). Teamwork refers to the ability of teams to coordinate and integrate the performance of multiple individuals (Salas et al., 2008). Taskwork refers to *what* teams are doing, whereas teamwork refers to *how* teams are doing it. For example, the task work of a Fire and Rescue team might involve extracting individuals from a burning building and supressing the spread of fire to prevent any further harm. To do this successfully, the team would need to engage in effective teamwork: the crew commander would need to communicate plans effectively, the operational staff members on the ground would need to coordinate their actions. Teamwork is therefore vital to enable team members to successfully engage in taskwork (Dinh & Salas, 2017).

In this thesis, I focus on identifying factors that support effective teamwork. Identifying what underpins effective teamwork is challenging as it is a complex, multi-faceted phenomena with many variables at play (Rosseau, Aube & Savoie, 2006). Prior research suggests it is the behaviours (e.g., communication), cognitions (e.g., shared mental models) and attitudes of team members (e.g., trust) interacting to enable the completion of shared team goals (Cannon-Bowers, Tannenbaum, Salas & Volpe, 1995; Salas, Sims & Klein, 2004; Salas, Sims & Burke, 2004). When a team is engaging in effective teamwork, they achieve more than individuals alone (Salas, Reyes & McDaniel, 2018; Stagle, Shawn, Burke & Pierce, 2006), reduce errors by providing back up and support to one another (Goodwin, Blacksmith & Coats, 2018), and complete joint-tasks more safely and efficiently (McEwan & Beauchamp, 2014).

Differentiating between extreme and conventional teams

Since the 1990s, researchers have increasingly considered how teamwork is influenced by the context in which teams are operating (*see* Hackman, 2002; Ilgen, 1999; Kozlowski & Ilgen, 2006; Kozlowski & Klein, 2000; Rousseau & Fried. 2001). Johns (2006) defines context on three levels; physical, task and social. *Physical context* refers to the physical attributes of the external environment (e.g., temperature, landscape). *Task context* refers to features of the task that are expected to influence behaviours, such as uncertainty and resource availability. *Social context* refers to the ties of team members, structure of relationships and the role of social influence on behaviours. More recently researchers have extended this definition to include the *temporal context*, referring to the aspects of the team and task environment in relation to time (Bell et al., 2018; Dinh & Salas, 2017; Mohammed, Hamilton & Lim, 2009). Context is a crucial aspect of understanding how teams operate as it can influence how team members interact with one another (Bell et al., 2018; Dinh & Salas, 2017; Hackman, 2002; Johns, 2006). Accordingly, a team that thrives in one context may not thrive in a different context.

There is a lack of empirical research that considers how context affects team behaviours (Dierdorff, Rubin & Morgenson, 2011; Goodwin et al., 2018). This is likely due to research on teams being carried out in a similar context of conventional teams (e.g., marketing teams, software development teams), with relatively little empirical research on teams in non-conventional settings (i.e., extreme teams, Bell et al., 2018; sports teams, McEwan & Beauchamp, 2014). This is problematic when considering the applicability of theoretical models and methods to measure and improve teamwork to non-conventional teams, since the majority have been developed largely with more conventional teams in mind (e.g., Wageman et al., 2005; Mathieu et al., 2008). In this thesis, I address this by focusing on advancing understanding of teamwork in ETs. Whilst, many aspects of teamworking, such as leadership (Burke, Shuffler & Wise, 2018) and communication (Uitdewilligen & Waller, 2018), have been shown to be important in ETs, much of the existing research on ETs remains largely theoretical, with an abundance of reviews and theoretical papers (e.g., Driskell, Salas & Driskell, 2018; Maynard et al., 2018; Roma & Bedwell, 2017). Given the limited empirical research on ETs and the often-complex structures and contexts in which they operate, more empirical research is needed to establish if teamwork operates in extremes as it does in other settings (Vessey & Landon, 2017).

Conventional teams and ETs differ on a number of key, inter-related features (*see* Figure 1). These relate to high consequences of poor performance, stress and uncertainty. I do not dispute that these features exist in conventional teams, nor do I suggest that findings in ETs have no applicability to conventional teams (*see* Hällgren, Rouleua & DeRond, 2018). It is, however, the extent to which they are present in extremes that differentiates the research context and illustrates the importance of this thesis examining ETs as their own distinct category. For example, whilst conventional teams may operate in highly pressurised organisations in which teamwork errors can have a large impact on financial success, ETs exist in contexts in which these errors can have catastrophic, even life or death consequences. For example, Risser et al. (1999) reviewed 54 incidents across eight different emergency departments in U.S hospitals. They found half of the deaths and permanent disabilities that occurred in these incidents could have been prevented had teamwork been better. In a review of the 1996 Mount Everest disaster in which eight climbers lost their lives, Kayes (2004) identified breakdowns in teamwork due to ill-defined goals and a maladaptive directive style of leadership likely contributed to the deaths.

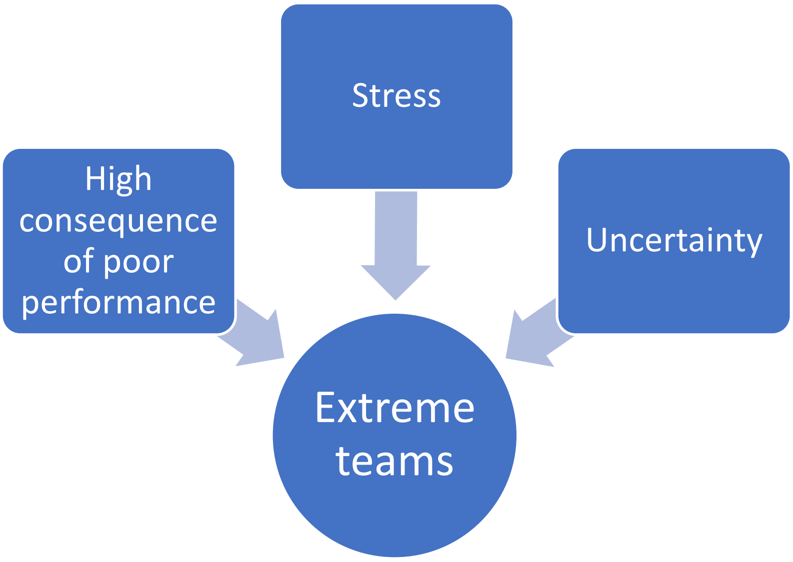


Figure 1. Defining characteristics of extreme teams

In addition to high consequences of poor performance, the second contextual difference between ETs and conventional teams is stress. Stress is defined as a process in which environmental demands create a perception that task demands exceed available resources, leading to undesirable physiological, psychological, behavioural or social outcomes (Salas, Driskell & Hughes, 1996). Whilst demands exist in conventional teams, they are likely to exist in a qualitatively different way (e.g., stress relating to production loss rather than the loss of life). The differences in contextual demands experienced in ETs may then drive the *type* of stress experienced in teams, which can amplify or alter the drivers of effective teamwork (Maynard et al., 2018; Driskell et al., 2018). Researchers have theorised that the experience of stress in extreme environments will impact behaviour in a number of ways and this is evident in empirical studies (Driskell et al., 2018). For example, stress has been found to reduce communication in aviation teams (Sexton & Helmreich, 2000), impair cognitive functioning in military teams (Wallenius, Larrson & Johannsson, 2004) and lead to a loss of team perspective in Navy Teams (Driskell, Salas & Johnston, 1999).

A further contextual difference between conventional teams and ETs is the degree to which team members experience uncertainty. Uncertainty exists in conventional teams (e.g., a lack of role clarity; Boynton, Gales & Blackburn, 1993; volatile business environments, Qi, Zhao & Sheu, 2011) and has been shown to impact performance (Geersbro & Ritter, 2010). However, in extremes, uncertainty is an inherent characteristic of the environment (Power & Alison, 2017a; Utidewilligen & Waller, 2018) and is therefore expected to be consistently present. ETs have been referred to in the literature as 'High Reliability Organisations'; differentiating them from conventional teams due to the need for these teams to operate safely and reliably for *sustained* periods in highly complex and uncertain environments (Klein, Bigley & Roberts, 1995; Roberts, 1990). Uncertainty in extremes can exist within the team (e.g., role uncertainty) and in the environment itself (e.g., dynamic changes in the environment) (Alison et al., 2015a). Uncertainty within the team can hinder coordination (Quarantelli, 1985; Weller, Janssen, Merry & Robinson, 2008), as well as reduce opportunities for collaboration due to a lack of clarity on responsibilities and capabilities (Power & Alison, 2017a). Uncertainty within the environment (e.g., relating to novel task requirements) increases the importance of good communication in teams and shared mental models (Marks, Mathieu & Zaccaro, 2000).

These contextual differences suggest the relative importance of variables may differ in ETs, when compared to conventional teams. For example, whilst extraversion has been associated with improved performance in conventional teams (Bell, 2007), there is some evidence that in isolated and extreme environments it can be detrimental (Palinkas, Gunderson, Holland, Miller and Johnson, 2000). This is thought to be due to isolated environments being less suited to individuals who benefit from diversified social interactions. Accordingly, it is important to conduct further empirical research on ETs to understand how knowledge from conventional teams transfers across to those working in challenging and extreme settings.

Intra-extreme team differences

Despite ETs sharing common characteristics such as high consequence of performance, stress and uncertainty, an important contribution to this thesis is to outline how there are marked differences between the challenges present in different *types* of extreme environments. The literature focuses either on multi-team systems (MTS), who form quickly in responses to crises (e.g., emergency response teams, disaster response teams, medical emergency teams) or on ETs who operate in isolated confined environments (ICE), in which teams must co-exist for longer periods in inhospitable climates (e.g., expedition teams, anti-poaching teams, teams in space). Whilst studies tend to acknowledge the unique features of MTS and teams in ICE, there is no formal typology within the literature.

In this thesis I suggest that developing a clearer distinction between types of ET would be useful for two key reasons: (i) recognising the salient challenges present in different types of extreme environment may better direct empirical research to the questions that need answering in each context, and (ii) distinguishing between different types of ET (i.e., MTS/ICE) would increase the applicability of findings and the utility of methods across different teams who share similar features (e.g., two different types of MTS both operating in an extreme environment). It should be noted that although the majority of ETs can be classified as either MTS or ICE, the two categories are not mutually exclusive. For example, in long distance spaceflight, teams on board the ship are isolated and confined (ICE), as well as being part of a larger network of teams (MTS) in maintaining communication with ground control (Anania, Disher, Kring, Iwig, Keebler, Lazzara & Salas, 2017). In the following section I define MTS and ICE, before highlighting why they differ and why future research might distinguish more clearly between the two.

Multi-team systems (MTS) comprise a network of component teams working together to achieve separate but related objectives within the framework of collective over-arching goals (Mathieu, Marks & Zaccaro, 2001; Shuffler, Jiménez-Rodríguez, & Kramer 2015). Examples of MTS operating in extreme environments include emergency responders (Waring, Alison, Shortland & Humann, 2019), aviation teams (Bienefeld & Grote, 2014), medical emergency teams (West et al., 2015) and disaster response teams (Rico, Hinsz, Burke & Salas, 2017). MTS are distinguished from other types of teams as they require collaboration *within* and *between* teams (Zaccaro, Marks & DeChurch, 2012). Component teams operating in a MTS may successfully manage within-team coordination but fail at between-team coordination (Firth, Hollenbeck, Miles, Ilgen & Barnes, 2015). The fact that MTS comprise multiple teams means that many of the theories and empirical evidence drawn from smaller, conventional work teams or large-scale organisations may not be generalisable to MTS (Lanaj, Hollenbeck, Ilgen, Barnes & Harmon, 2013). This is because MTS are larger and more specialised than conventional work teams (e.g., they tend to be brought together to deal with emergencies) but are smaller and less formal in structure than large organisations (Lanaj et al., 2013).

MTS also form quickly in response to emergency incidents. This creates unique challenges in managing conflicting objectives across component teams, a lack of familiarity amongst component teams and a need to maintain communication channels across larger, often disparate networks (Critchton, Flin & Rattray, 2000; Fodor & Flestea, 2016; Lacerenza, Rico, Salas & Shuffler, 2014; Shuffler & Carter 2018). Whilst the science of MTS remains in the early stages, in line with the development of more sophisticated theoretical models, there is a need to conduct empirical research to explore how these models work in practice (Landon, Slack & Barrett, 2018; Luciano, DeChurch & Mathieu, 2018; Shuffler & Carter, 2018; Power, 2018).

In contrast to MTS, some ETs operate in isolated and confined environments (ICEs) for prolonged periods. This further exacerbates the difficulty of applying theoretical models of teamwork that have been developed with conventional teams as many are based on data from cross-sectional studies (Golden, Chang & Kozlowski, 2018). Examples of ICE include polar exploration, mountaineers, long-distance space flight, anti-poaching teams, long-distance sailing and submariners (Palinkas, 2003). Teams in ICE are exposed to a number of physical and psychological stressors, often in inhospitable climates and with little contact with the outside world (Barrett & Martin, 2014). The physical isolation of teams in these environments and the confinement with a small number of individuals for prolonged periods suggests different types of challenges compared to those faced by MTS. For example, the characteristics of many environments in ICE (e.g., weather conditions, difficult terrain) will likely increase the extent to which the physical context influences team member dynamics (Dinh & Salas, 2017). In MTS, the main challenges instead relate to bringing together previously unacquainted teams to complete complex tasks, that require coordination *across* teams (Shuffler et al., 2015). Thus, in this thesis I focus on the development of effective team processes in MTS and how this might be expedited by familiarity between team members (*see Chapter III* and *IV*). I do this by researching emergency responders as an example of a MTS.

The prolonged length of time that individuals work together in ICE also differentiates them from the sometimes-brief period with which MTS may work together. This creates challenges in balancing the personalities and values of multiple individuals over time (Roma & Bedwell, 2017), managing changes in the physical environment (Barret & Martin, 2014), in addition to feelings of boredom and monotony within the group (Leon, Kanfer, Hoffman & Dupre, 1994). The current literature and theoretical frameworks relating to team effectiveness do not provide conclusive, evidence-based recommendations to support team performance in isolation for prolonged periods, in the face of persistent stressors (Golden et al., 2018). Much of the research on ICE has so far considered the individual characteristics that are best suited for survival in this context (e.g., coping strategies [Smith, Barrett & Sandal, 2018]; personality traits [Sandal, Endresen,Vaerners & Ursin, 1999]), however there is a growing emphasis on the need to consider team variables such as cohesion and how they emerge over time (Roma & Bedwell, 2017; Vessey & Landon, 2017, *see Chapter V*). This is because when teams are in ICE, they not only work with one another but also live with one another. The continued time spent together in isolation increases the likelihood of fault lines within the team (e.g., differences in personalities/values) leading to conflict (Kealey, 2004). In such instances, maintaining a high level of cohesion will be especially critical to team performance (Stuster, 2011). This is explored in *Chapter V* of this thesis by monitoring how cohesion emerges and is sustained in ICE over time. I do this by focusing my research on expedition teams.

Research Questions

Identifying the different challenges faced by MTS and teams in ICE, emphasises why it is important to distinguish between *types* of ETs (*see* Figure 2). Further, it highlights how different aspects of teamworking may matter more, dependent on the whether an ET is defined as MTS or operating in ICE. This may help to direct empirical research towards the most appropriate questions that need answering in each context. For example, whilst cohesion is cited as being especially important for ETs in ICE (Stuster, 2011), this may not be the case for MTS. Research suggests that cohesion in component teams in MTS can actually increase inter-group competition and undermine overall system performance (DiRosa, 2013). Moreover, researchers have argued the fast forming and quickly disbanding nature of some MTS (i.e., emergency response teams) can mean variables like cohesion which emerge and fluctuate over time will be less likely to have a meaningful impact on performance (Power, 2018).

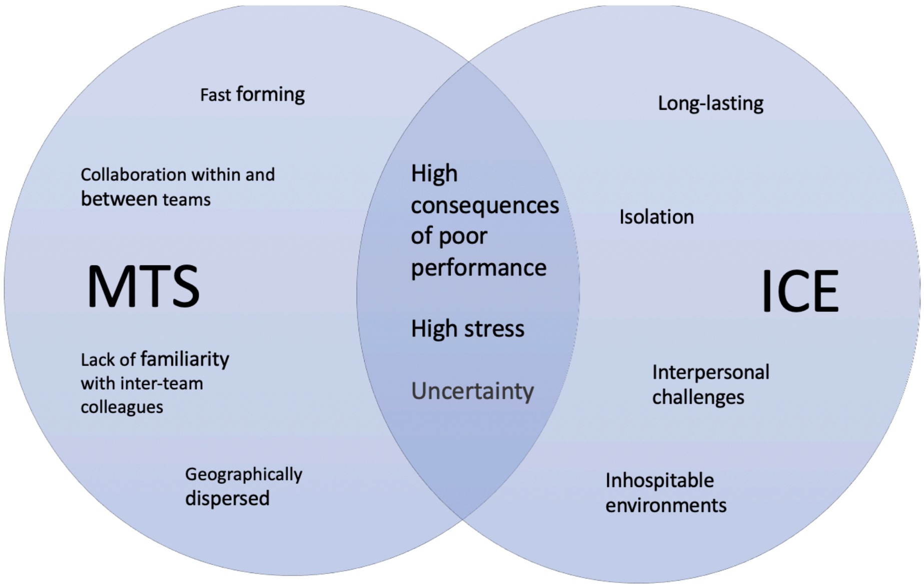


Figure 2. Typology of intra-extreme differences

Greater clarity within the literature on the different types of ET will also ensure findings from one type of team (e.g., emergency response teams) can be applied to teams of the same characteristics (i.e., other MTS such as disaster response teams). For example, researchers have suggested the evidence base of expedition teams in ICE can be used to inform the understanding of other teams in isolated settings such as defence and security personnel involved in counter-terrorism operations in remote and inhospitable climates (Smith & Barrett, 2019). In light of this, the first aim of this thesis was to identify factors that support effective teamworking in different types of extreme environments.

RQ1: *What factors support effective teamwork in different types of extreme environments (MTS and ICE)?*

To understand the factors that support effective teamwork in MTS and ICE, this thesis also considered the methods and analytical approaches suited to studying different types of ET. Bell et al. (2018) suggest researching ETs is challenging as they are difficult to access. Moreover, small sample sizes challenge the utility of many traditional statistical techniques and methods of data collection. For example, the accessibility of conventional teams enables researchers to conduct large scale, longitudinal studies in which complex statistical modelling of big data sets can be used to further understanding (e.g., Mathieu, Kukenberger, D'Innocenzo & Reily, 2015). Whilst research in ETs has commonly used methods such as interviews and questionnaires to provide a descriptive overview of the challenges present in extreme environments (e.g., Gillespie, Gwinner, Chaboyer & Fairweather, 2013; Power & Alison 2017a; Wauben et al., 2011), these methods are not able to make direct observations of behaviours or to manipulate specific variables to empirically test theoretical models.

When studying emergency response teams, immersive simulations offer an alternative approach to questionnaire and observational studies (*Chapter II*), allowing researchers to re-create the stressors present in extremes, whilst conducting rigorous empirical research (Manser, Dieckmann Wehmer & Rall, 2007). Simulations allow researchers to collect behavioural data that is needed to understand teamwork in this context (Alison, van den Heuvel, Waring, Power, Long, O'Hara & Crego, 2013; Rosen et al., 2008), whilst employing mixed methods analyses to better understand the data (*Chapter III, IV*). In conventional teams, behavioural data may be collected using a variety of *in-situ* methods (e.g., observational studies, ethnography). However, during the response to an emergency this might not be possible due to the security sensitive nature of the context (e.g., an *in-situ* observation of the response to a terrorist incident) and the possibility that having researchers present would jeopardise the response.

Simulations have also been used to study teams in ICE (e.g., those in long distance spaceflight, *see* Landon et al., 2018). However, they are unlikely to be appropriate when studying expeditions teams and for other mobile isolated teams (e.g., anti-poaching teams). This is because a simulated environment is unlikely to reflect the inhospitable climate (e.g., weather conditions) and arduous physical challenges of being on expedition. Demonstrating the importance of selecting methodologies and analytical techniques that are appropriate to the type of team studied is an important contribution of this thesis. Given that the remoteness of expedition teams excludes the use of wearable devices (e.g., sociometric badges), collecting self-report data 'in the wild' will be key to advancing understanding of how team members interact and work effectively (Kozlowski, 2015). In studying expedition teams, I identify the benefits to using a longitudinal diary methodology (*Chapter V)* to monitor emerging team dynamics over time (Kjaegaard et al., 2015; Smith et al., 2018). Intensive longitudinal designs ensure that even with relatively low sample sizes, day-to-day fluctuations are captured, improving reliability and increasing the likelihood of inferring causality (Golden et al., 2018).

RQ2: *What methods and analytical approaches are suitable for studying teams in different types of extreme environments?*

Background to the research

This Chapter has thus far provided an overview of the thesis, and the two main research questions that will be addressed. To address these questions, it is necessary to first review teamwork models. This is because there exist many models and approaches to studying teams (e.g., Ilgen, Hollenbeck, Johnson & Jundt, 2005; Klein & Kozlowski, 2000; Kozlowsi & Ilgen, 2006; McGrath, 1964; Marks, Mathieu & Zaccaro, 2001; Millward, Banks & Riga, 2011) and it is important to be clear how teams are conceptualised in this thesis. Second, it is necessary to review the existing research on emergency response teams and expedition teams to outline the variables focused on in subsequent empirical Chapters. This is because teamwork is a multi-faceted, complex phenomenon and considering all aspects of teamworking (e.g., communication, coordination, leadership, trust, individual differences, team size, familiarity) is beyond the scope of the thesis. Accordingly, this section identifies variables that are appropriate for further empirical study. This is done by: (i) identifying variables that are relevant to the contextual challenges present in each type of extreme environment (MTS and ICE); (ii) identifying variables that are hypothesised to improve performance in ETs, but have not been tested empirically in this context (familiarity in *Chapter IV,* cohesion in *Chapter V)* and (iii) identifying variables that have been studied in ETs but due to recent theoretical advancements would benefit from further study (communication and coordination in *Chapter III)*.

Models of teamwork

The Input-Process-Outcome (I-P-O) model, first proposed by McGrath (1964) and later expanded by Hackman (1986), provided the foundational understanding to much of the team literature (McEwan & Beauchamp, 2014; Mathieu et al., 2018; Mathieu et al., 2008; Kozlowski & Chao, 2018; Salas et al., 1992). According to this model, inputs are defined as the antecedent factors that influence team member interactions and thus outcomes (Mathieu et al., 2008). **Inputs** include the composition of the team (e.g., personality, team size, familiarity of team members) task characteristics (e.g., demand, urgency) and the resources available to the team (e.g., information) (Kozlowski & Ilgen, 2006). **Processes** (e.g., communication, coordination, cooperation) refer to how team members interact with one another to complete tasks (Kozlowski & Chao, 2018), representing the coordinating mechanisms and behaviours used by teams to achieve collective goals (Power, 2018). Finally, **outcomes** are understood according to Hackman's (1986) typology of: (i) performance of the team; (ii) team member satisfaction; and (iii) team viability to sustain effective working. Within the I-P-O framework, the relationship between inputs and outcomes is mediated by processes.

The original formulation of the I-P-O model has been criticised on several grounds. Firstly, researchers noted that it failed to provide a clear definition of the term processes, which was deemed problematic given the diversity of the term (Marks et al., 2001; Millward et al., 2010). For example, whilst processes refer to how team members interact with one another within a task environment, the term had been used to describe a broader range of variables such as cohesion and collective efficacy. Marks et al. noted that many mediational factors (such as cohesion, collective efficacy) that influence team outputs were not processes but were better defined as "emergent states". Emergent states refer to the "properties of the team that are typically dynamic in nature and vary as a function of team context, input, processes and outcomes" (Marks et al., 2001, p. 357). Broadly, researchers have identified three types of emergent states: (i) cognitive (e.g., shared mental models, *see* Mohammed & Dumville, 2001), (ii) affective (e.g., cohesion, *see* Molleman, 2005), and (iii) motivational (e.g., collective efficacy, *see* Peterson, Mitchell, Thompson & Burr, 2000).

Secondly, researchers noted that the original I-P-O model was too static to adequately represent the complexities of team behaviour, due it being developed largely from studies with cross-sectional designs (Cronin, Weingart & Todorova, 2011; Marks et al., 2001). Teamwork research has advanced such that it acknowledges that teams exist within a wider environment (Kozlowski & Ilgen, 2006), adapting over time and in response to environmental demands (Hackman, 2003; Kozlowski & Ilgen, 2006; McGrath & Berdahl, 2002). For example, inputs, processes and outcomes are expected to interact over time, both in the external environment and within the environment of the team, influencing team behaviour in a reciprocal manner (Kozlowski, Gully, Nason & Smith, 1999). Similarly, the I-P-O model implies that teams progress in a linear manner, moving from the input to the process to the outcome. However, team behaviours are dynamic with empirical research finding relationships between inputs and emergent states (e.g., personality and cohesion, *see* Barrick, Stewart, Neubert & Mount, 1998) and between processes and processes (e.g., De Drue & Weingart, 2003).

In light of these theoretical advancements, Ilgen et al. (2005) expanded the I-P-O framework to acknowledge the role of emergent states and the dynamic nature of team interactions, introducing the input-mediator-outcome-input (IMOI) framework. According to this model, the relationship between inputs and outcomes occurs in a cyclical process, influenced by mediators (processes and emergent states). The IMOI framework is used widely in the research of teams across a range of contexts (Golden et al., 2018; Jaca, Viles, Tanco, Mate & Santos, 2013; McEwan & Beauchamp, 2014; Zhou & Wang, 2010) and is considered a seminal piece of theoretical research in this field (*see* Kozlowski & Chao, 2018; Mathieu et al., 2008; Salas et al., 2008; Salas, Rico & Passmore, 2017). Consistent with other research of ETs (e.g., Golden et al., 2018), I use the IMOI framework as the conceptual foundation to understanding teamwork in this thesis. As both processes and emergent states are studied in subsequent empirical Chapters, it was important to adopt a framework that clearly distinguishes between the two. In addition, the IMOI framework more clearly recognises the role of time in team-based interactions than the original IPO framework. This is important for *Chapter III* and *V* where teamwork is studied over time.

Multi-team Systems

Multi-team systems (MTS) are defined as a network of component teams working together to achieve separate but related objectives within the framework of collective over-arching goals (Mathieu et al., 2001; Shuffler et al., 2015). In this thesis, I study emergency response teams as an example of a MTS. In the U.K, an emergency is defined as any event that threatens the welfare, environment or security of the U.K, including acts of terrorism or war (Civil Contingencies Act, 2004). A "major incident" is a specific type of emergency requiring the coordinated response of multiple agencies, with the potential for mass casualties and heightened media interest (Cabinet Office, 2012; *Chapter III* for a more detailed overview). Examples of major incidents include large scale traffic accidents, terrorist attacks and wide-spread flooding.

Major incidents are highly complex and ambiguous contexts that challenge decision-making and teamwork (Alison & Crego, 2008; McMaster & Baber, 2012).

These characteristics place considerable psychological, physiological and interpersonal demands on team members, defining major incident response as an extreme environment (Manzey & Lorenz, 1998; Power, 2015). Responders must form *ad hoc* teams, pool a variety of expertise, develop role specialisation and achieve multiple objectives in parallel, balancing the needs of intra-agency and inter-agency goals (Crichton et al., 2000). For example, whilst all responders will be seeking to sustain life and reduce harm to the public, agencies will also be tasked with additional, differing objectives. A police commander might be tasked with preserving evidence from the scene and collecting information from witnesses, whereas an ambulance commander might be tasked with prioritising casualties. It is important that operational team members on the ground and strategic commanders providing oversight, can achieve these objectives in parallel.

Despite many studies exploring teamwork in emergency response, recent government reports have identified continued challenges (e.g., Kerslake, 2018; Pollock, 2017). This is mirrored by findings in the literature. For example, Alison et al. (2015a) identified that 75% of uncertainties that commanders reported experiencing during response could be attributed to the team (e.g., poor role understanding), with only 25% attributed to the task environment itself (e.g., urgency of situation). Using retrospective analysis of communications during an emergency incident, Fodor and Flestea (2016) identified breakdowns in communication across different agencies delayed the emergence of situational awareness, leading to a reduced understanding of the task in the early stages of response. These findings highlight the important practical applications of further research in this field, as team-based failures continue to create challenges.

Most research on emergency response teams focuses on how team processes facilitate effective response (e.g., communication across different agencies, *see* Fodor & Flestea, 2016; Waring et al., 2019). It is, however, important to acknowledge relevant input factors. MTS are particularly suited to responding to dynamic, complex emergency incidents because they can utilise the diverse skills and abilities of different component teams (Marks, DeChurch, Mathieu, Panzer & Alonso, 2005; Zaccaro, Marks & DeChurch, 2012). Whilst the original IMOI (Ilgen et al., 2005) does not reference diversity *across* teams as an input factor, Zaccaro et al. (2012) extend the IMOI framework to refer to the "compositional attributes of the MTS" (i.e., which component teams form the MTS). More complex environments (such as emergency response) are said to require a higher level of functional diversity and thus a greater number of component teams (James & Wooten, 2010; Mathieu, Marks & Zaccaro, 2001).

Diversity can nevertheless intensify differences between component teams and increase the challenge of managing team processes (e.g., coordination) across the network (Luciano et al., 2018). Research from emergency response teams suggests that despite needing diversity to ensure an effective response, it can cause problems with communication and lead to reduced situation awareness (Fodor & Flestea, 2016). Relatedly, functionally diverse MTS will naturally tend be larger than less functionally diverse MTS as they will include a larger number of different agencies. Research suggests that an increase in MTS size makes managing team mediators more difficult (Lanaj et al., 2013). This is supported by evidence in the conventional team literature (LePine, Piccolo, Jackson, Mathieu & Saul, 2008) and from an empirical study of multi-agency emergency responders (Alison, Power, van den Heuvel, Humann, Palasinski & Crego, 2015b). The current perspective would suggest the size and functional diversity of a MTS may be a double-edged sword in terms of predicting team outcomes (Shuffler et al., 2015). Too small and the system may not successfully achieve the MTS goal, too large and this may lead to difficulties in establishing shared understanding and effective coordination (Burke, Shuffler, Heyne, Salas, & Ruark, 2014; Weaver, Dy & Rosen, 2014)

This thesis examines how familiarity may help to overcome some of the challenges of working within a functionally diverse MTS (*see Chapter IV).* Evidence from conventional teams shows when members work together for a long period of time, they become familiar with the task and each other, which has been found to benefit team performance across a range of contexts (Huckman et al., 2009; Joshi, Hernandez, Martinez, Abdel-Fattah & Garden, 2017; Katz, 1982). The positive effect of familiarity on team performance is thought to be underpinned by better developed transactive memory systems in familiar teams (Austin, 2003), and team mental models (Salas et al., 2005). Transactive memory systems are defined as the knowledge possessed by different team members and awareness of who knows what, and team mental models are defined as the shared knowledge of team functioning and expected behaviours. The effect familiarity on teamwork in MTS has yet to be tested empirically (*see Chapter IV)*, although research suggests that increased familiarity can provide a greater understanding of one another's roles and thus improve the timely and appropriate exchange of information (Waring, Alison, Carter, Barrett-Pink, Humann, Swan & Zilinsky, 2018). Thus, whilst the role of familiarity within the MTS literature remains under-researched (Turner et al., 2019), it has been hypothesised that higher levels of stability across component teams will expedite team processes and improve efficacy (Shuffler et al., 2015).

As stated, most research on emergency response teams has focused on understanding how team processes support effective response in crises. Unlike conventional work teams or teams operating in ICE, emergency response teams tend to be fast forming, with team members who may rarely work together, leading to improvised organisational structures (Kapucu, 2006; Shuffler et al., 2015). This can mean that emergent states such as cohesion which develop over time may not be as important as team members are only working together for short periods (*see* Mathieu et al., 2001). It is therefore more important in this context that researchers focus on identifying solutions to improving team processes (Power, 2018).

Specifically, research has focused on two processes; coordination and communication, referred to in the literature as the cornerstones to effective emergency management (Haddow & Bullock, 2003; Helsloot, 2005; Junglas & Ives, 2007; Power, 2018; Smith & Dowell, 2000; Utidewilligen & Waller, 2018). Coordination is defined as the enactment of behavioural and cognitive mechanisms that enable team members to synchronise their efforts to achieve goal related outcomes (De Dreu & Weingart, 2003; Marks et al., 2001). Communication is defined as a reciprocal process, involving the sending and receiving of information by team members, shaping team attitudes, behaviours and cognition (LePine et al., 2008).

In this thesis, I continue to focus on communication and coordination processes in MTS as they are vital to address the challenges of working with multiple agencies in crises. Coordination enables team members to work together, combining the skills and knowledge of multiple agencies (Power, 2018). Failure to coordinate effectively during an incident can lead to a failure in team action and delays in decision-making (Alison et al., 2015b) and is associated with other related aspects of teamworking. For example, in an interview study with 31 commanders from the emergency services, Power and

Alison (2017a) identified that poor role understanding contributed towards poor coordination. Without a clear understanding of roles and responsibility, teams tend to operate independently, failing to act collectively as a whole (Salmon, Stanton, Jenkins & Walker, 2011; Perry & Wears, 2011). Communication during emergencies allows disparate team members who may not be co-located to develop a common operational understanding of the incident and develop shared situational awareness (Salas, Prince, Baker & Shrestha, 1995). In a simulation study of emergency response teams, higher performing teams spent more time structuring and sharing information than lower performing teams (Utidewilligen & Waller, 2018).

Existing research has identified the importance of communication and coordination in emergency response teams, concurrent with findings in the conventional team literature (*see* Haddow & Bullock, 2003; Hoegl, Weinkauf & Gemuenden, 2004; Marlow, Lacerernza, Paoletti, Burke & Salas, 2018). In this thesis, I make three important contributions to existing research by: (i) approaching emergency response teams from a MTS perspective; (ii) exploring communication and coordination over time, and (iii) exploring how communication and coordination differ according to the level of familiarity in teams.

The first contribution to existing research on communication and coordination is to approach the study of emergency response teams from a MTS perspective. Despite researchers acknowledging emergency response teams as multi-agency teams, many studies do not reference theoretical models and frameworks of MTS and instead focus on the practical implications of findings (e.g., Chen, Sharman, Chakvarati, Rao & Upadhyaya, 2008; Mishra, Allen & Pearman, 2011; Salmon et al., 2011). Practical implications are important to reduce the number of team-based errors in this setting, however failing to acknowledge the MTS literature may affect the interpretation of findings. This is because processes are expected to occur differently in MTS (*see* Mathieu et al., 2001; Zaccaro et al., 2012). For example, researchers have debated the benefits of monitoring frequency and patterns of communication as markers of performance in ETs, as lower levels of information sharing can be indicative of greater implicit knowledge of one another's working (*see* Stout, Cannon-Bowers, Salas & Millanovich, 1999; Stachowski, Kaplan & Waller, 2009). However, this seems to be more appropriate for smaller, more compact teams in crisis (e.g., in a command and control room, *see* Waller, Gupta & Giambatista, 2004). In Luciano et al.'s (2018) meso-theory of MTS functioning, low levels of *information opacity* (the absence and ambiguity of information across teams) facilitates a shared understanding of the situation, which in turn, leads to better performance. MTS are often disparate in nature, involving team members who may never have worked together before, thus to maintain shared awareness, information must be continually shared and updated as component teams will have access to different streams of information (Fodor & Flestea, 2016; Mishra et al., 2011; Waller & Uitdewilligen, 2008).

Approaching the study of emergency response teams from a MTS perspective may shed new light on our understanding of communication and coordination in this setting. This is explored in *Chapter III* and *IV* by coding interactions in emergency response teams to map out communication networks and identify verbal indicators of coordinating behaviours. Doing so has important implications to the wider MTS literature. Whilst the theoretical literature on MTS has advanced considerably in recent years, empirical studies lag behind and more research is needed to begin understanding how frameworks and models work in *in situ* (Shuffler et al., 2015). For instance, whilst detailed theoretical frameworks of coordination in MTS exist, there is a lack of empirical research identifying the behaviours that underpin it in practice (Rico, Hinsz, Burke & Salas, 2017; Wijnmaalen, Voordijk & Rietjens, 2018). Further, this approach will ensure findings have practical implications beyond the specific context of investigation. For example, whilst errors have been attributed to inter-team working in emergency responders, (*see* Kerslake, 2018; Waring et al., 2019), similar problems arise in other multi-agency teams (i.e., medical teams, Gerber et al., 2016; aviation, Bienefeld & Grote, 2014; military teams, Wijnmaalen et al., 2018).

The second contribution to existing research is to study communication and coordination across time. There is a lack of empirical studies of MTS and emergency response teams that consider the role of time in shaping team processes (Bienefeld & Grote, 2014; DeChurch & Marks, 2006). This is problematic as MTS are expected to continually adjust their behaviours in response to changing demands, suggesting we can expect to see differences in which behaviours contribute to effective performance at different points in time (Aiken & Hanges, 2012; Waring et al., 2019). During different phases of emergency response, we can expect certain component teams (focal teams) to be more important in the attainment of over-arching MTS objectives (Davison, Hollenbeck, Barnes, Sleesman & Ilgen, 2012). In their Meso-Theory of MTS functioning, Luciano et al. (2018) refer to this as the *fluidity of the structural configuration* and describe how we can expect to see changes in how team members interact with one another across time, dependent on task demands. For example, it might be that centralised communication structures (i.e., those that rely on a single focal team) are especially problematic in the earlier phases of an emergency as they can delay the development of shared situational awareness. To generate a better understanding of how this works in practice, it is necessary to study temporal changes in communication and coordination. The importance of measuring team processes over time is discussed in detail in *Chapter II,* and is explored empirically in *Chapter III,* by modelling how communication networks and coordinating behaviours in MTS change during different phases of emergence response.

As discussed previously, there is reason to believe that increasing familiarity across the MTS will expedite team processes and improve performance (Waring et al., 2018; Lucianio et al., 2018). This is supported by findings in the conventional team literature suggesting familiarity influences team behaviours and improves performance outcomes (*see* Espinosa, Slaughter, Kraut & Herbsleb, 2007; Harrison, Mohammed, McGrath, Florey & Vanderstoep, 2003). Accordingly, the final contribution to existing research is to explore if communication and coordination improve when team members are familiar with one another. There are no empirical studies on the role of familiarity in MTS and this is addressed *Chapter IV.*

Teams in Isolated and Confined Environments

Some ETs operate in isolated and confined environments (ICE). These environments are characterised by: (i) isolation from family and friends; (ii) confinement/restrictions on usual mobility patterns and (iii) inhospitable and hazardous climates (Golden et al, 2018; Nicolas, Seudfeld, Weiss & Gaudino, 2016). The type of team studied in ICE in this thesis is expedition teams. Whilst there are other teams that may be more readily identifiable as ICE (e.g., teams in long distance spaceflight, submariners), I chose to study expedition teams because they offer an accessible alternative with which to generate a greater understanding of teamwork in this context (Golden et al., 2018).

An expedition is defined as a purposeful journey, undertaken for reasons of adventure, exploration and scientific discovery (Johnson, Anderson, Dallimore, Winser & Warrel, 2008). Expeditions can vary in length and level of risk. For example, an expedition may be a week-long trek on a known mountain trail or a lengthier journey in an uncharted cave system. Expedition environments are characterised by contextual (i.e., relating to the physical context, such as weather conditions, inhospitable climates), psychological (i.e., relating to the task context, such as monotony, exhaustion) and interpersonal challenges (i.e., relating to the social context, such as lack of personal space, limited contact with family and friends) thus defining them as an extreme environment (Barrett & Martin, 2014; Smith et al., 2018). Moreover, when on expedition, teams are often physically isolated and confined, spending large parts of the day in tents due to poor weather conditions (Palinkas & Suedfeld, 2008). This further defines them as ICE. The isolation, in combination with the arduous physical challenges on expedition can make establishing and maintaining effective teamwork difficult, exaggerating individual differences between team members and increasing the likelihood of interpersonal conflicts (Palinkas & Seudfeld, 2008; Stuster, 2011).

Despite the challenges to working as a team on expedition, teamwork is vital for success. Team members in ICE must coordinate and communicate with one another to cope with the challenges of the environment and reduce risks to safety (Bishop, Morphew & Kring, 2000; Driskell et al., 2018). A qualitative study of an Arctic crossing illustrates the importance of teamwork in this setting, with one team member noting: "If you don't have the team you have nothing. Have team members who have social intelligence… anyone can learn tasks" (Leon, Sandal, Fink & Ciofani, 2011, p.14). In a similarly isolating environment at an Antarctic station, a diary study of 32 individuals revealed that one third of diary entries were about interpersonal relationships (e.g., conflict amongst team members) (Weiss & Gaud, 2004). The authors concluded that social relationships were important for successful adaptation to the environment. This literature shows why it is important to study teamwork in ICE and demonstrates that, whilst technical skills and abilities are important, there is a reliance on the interpersonal relationships between team members (Ortner, 1999)

Within the literature, researchers have tended to focus on individual adaptability and performance in expedition teams, leading to a focus on stress, individual coping mechanisms and health/well-being (Kahn & Leon, 1994; Smith, Kinnafick & Saunders, 2017; Smith et al., 2018). In a systematic review of adaptability in ICE, intelligence, emotional stability, conscientiousness, introversion and task- oriented coping mechanisms were identified as important factors for performance (Bartone, Krueger & Bartone, 2018). Despite a recognition of the relationship between these variables and teamworking (Kahn & Leon, 1994; Leon et al., 2011), there has been limited empirical research that has focused on what factors contribute to effective teamwork in this setting. Of this research, there has been a focus on input factors such as the composition of the team and the ways in which gender, personality and values might allow team members to thrive in ICE and achieve team goals (Palinkas & Suedfeld, 2008). For example, research has found higher levels of competitiveness in all male expedition teams (Bishop, Grobler & SchjØll, 2001), with some evidence that introducing female team members can normalise group behaviours and promote individual and team performance (Leon, 1991). Research in an analogous context to an expedition environment of individuals overwintering at an Antarctic station identified low levels of extraversion as a significant predictor of performance (Palinkas, Gunderson, Holland, Miller & Johnson, 2000). Notwithstanding the importance of the composition of teams on expedition, there is also a need to advance this research into understanding the ways in which other factors of teamwork operate (*see Chapter V)*.

There is a growing recognition of the importance of team processes on expedition (communication, coordination etc., *see* Barrett & Martin, 2014), with empirical research to suggest that honest and open communication in teams can improve performance (Kjærgaard et al., 2015; Smith et al., 2017). In addition to team processes, team cohesion has been identified as crucial for performance in ETs (Landon & Vessey, 2017), with researchers suggesting that this will be especially so in ICE (Landon, Vessey & Barrett, 2015; Roma & Bedwell, 2017). Stuster (2011) suggests this is because the environment amplifies the relationship between cohesion and performance. With team members in relative isolation, this creates more opportunities for breakdowns in cohesion to occur which can subsequently impact social relationships and task completion. A special consideration of how cohesion is established and maintained in ICE is therefore necessary (Vessey & Landon, 2017). Whilst acknowledging the importance of team composition and processes, I focus here on how team cohesion is developed and sustained in an expedition environment (*see Chapter V*). Cohesion is defined as the shared tendency for the team to remain united in achieving a common goal (Casey-Campbell & Martens, 2009). It has also been referred to as the commitment of a team to work on a task constructively, whilst maintaining social relationships (McClurg, Chen, Petruzelli & Thayer, 2017).

Team cohesion is well-researched in conventional teams and has been identified as an essential component of team performance (Beal, Burke & McLendon, 2003; Evans & Dion, 2012; Salas, Grossman, Hughes & Coultas, 2015). High levels of cohesion can facilitate decision-making under pressure (Zaccaro, Gualtieri & Minnionis, 1995) and improve team member attitudes (Greer, 2012) and adaptability (Maynard & Kennedy, 2016). Historically, researchers have debated the dimensionality of cohesion (Festinger & Thibaut, 1951; Seashore, 1954), however there has been a consensus of recent that it is best understood in terms of its task and social components (Boyd, Kim, Ensari & Yin, 2014; Carless & De Paola, 2000; Salas et al., 2015). Task cohesion is defined as the shared commitment to the task, and social cohesion as the interpersonal bonds existing between team members (Mikalachki, 1969). The fact that cohesion encompasses task focus and social relations makes it especially relevant to studying teams in ICE, who by necessity work and live alongside one another (Vessey & Landon, 2017).

Cohesion is defined as an emergent state within the IMOI framework and is therefore expected to change over time (Carless de Paola, 2000; Ilgen et al., 2005). Accordingly, to understand the relationship between cohesion and performance, it is important that measures are taken longitudinally (Braun, Kozlowski, Brown & DeShon, 2020; Greer, 2012; McClurg et al., 2017). This is reflective of calls within the broader team literature to consider temporal influences on teamwork, indicating that there are applied and theoretical reasons to focus on cohesion in expedition teams (Kozlowski & Ilgen, 2006; Salas et al., 2018). Focusing on cohesion in expedition teams offers an opportunity to collect ecologically valid data over time, which can be used to inform understanding of other teams working in analogous settings (e.g., long distance spaceflight) (Sandal, Smith & Leon, 2017). In an extreme and challenging environment, where contextual factors (e.g., weather conditions) can vary considerably, micro-variations in cohesion are expected to influence performance. Limited evidence from conventional teams suggests that we can expect to see a positive influence of cohesion on perceived performance over time (*see* Mathieu et al., 2015; Braun, Kozlowski, Brown & DeShon, 2020). There is, however, little evidence of how cohesion is associated with performance in ICE and this is addressed in *Chapter V*.

Empirical studies of cohesion in expedition teams tend not to focus on cohesion as the main topic of investigation, as such it is often measured with a single item. A study of a team completing a 61-day trek through parts of Alaska identified positive associations of cohesion with communication, perception of fairness and the perceived quality of the leaders' decisions (Leon et al, 1994). Other findings suggest that cohesion is related to role hierarchy and acceptance of leadership on expedition (Eys, Ritchie, Little, Slade & Oddson, 2008), as well as changes in emotion and mood (Wagstaff & Weston, 2014). Research has also shown that the composition of teams may affect the way in which cohesion emerges with evidence of higher cohesion in mixed gender teams (Bishop, 2004) and in teams with compatible values (Sandal, Vaernes & Ursin, 1995). Only a handful of studies have monitored cohesion and performance over time in this setting. One diary-based study in the Antarctic identified a reciprocal relationship between the two over time (Kozlowski, Biswas & Chang, 2015). A further study identified a relationship between cohesion and task difficulty in an all-female climbing team, with cohesion increasing steadily up until the point the group engaged in the most challenging task (Allison, Duda & Beuter, 1991). Cohesion then began to tail off at the end of the expedition, coinciding with an increase in group conflict.

Despite numerous studies demonstrating the criticality of cohesion in ICE (Stuster, 2011; Wood, Schmidt, Lugg, Ayton, Phillips, & Shepanek, 2005), few studies have focused specifically on measuring cohesion in this context and how it evolves over time. Moreover, none of the existing studies use validated measures of cohesion or take consistent daily measures that can account for micro-variations. *Chapter V* addresses these limitations and directly answers the call for more research in this setting that seeks to understand how cohesion emerges and is sustained over time in ETs (Landon et al., 2015; Vessey & Landon, 2017).

In addition, this thesis will explore how the personality traits of team members shape the development of cohesion. Existing reviews and meta-analyses suggest we can expect to see a relationship between team composition and cohesion (Bell, 2007; Bell

& Outland, 2017). Studies of personality composition and cohesion in conventional teams have identified positive associations between emotional stability and cohesion (Barrick et al., 1998) and agreeableness and cohesion (Bradley, Baur, Banford & Postlethwaite, 2013). In examining cohesion and team composition over time, one study found teams who are highly agreeable tend to exhibit higher levels of cohesion over time (Acton, Braun & Foti, 2019). There is no empirical research to my knowledge that has conducted similar research on teams in ICE and this is addressed in *Chapter V.*

Conclusion

Reviewing the literature on emergency response and expedition teams has identified factors central to understanding teamwork in these contexts and these factors are explored in subsequent empirical Chapters. Emergency response teams are consistently referenced as an example of a MTS in the literature with associated challenges often acknowledged (e.g., Kozlowski & Chao, 2018; Luciano et al., 2018; Shuffler et al., 2015), yet theoretical models and frameworks of MTS are not adopted in empirical studies. By researching emergency response teams from a MTS perspective, findings will generate a better understanding of team processes in this context and contribute to a broader understanding of factors that support teamwork in multi-agency teams. Further, this thesis extends prior research by exploring how team processes change across time in MTS (*Chapter III)* and if they are affected by familiarity amongst team members (*Chapter IV)*. In studying expedition teams as an example of a team in ICE, findings in this thesis build on existing evidence that acknowledges the unique challenges that teams in this context face (Golden et al., 2018). By measuring cohesion over time in expedition teams (*Chapter V)*, this thesis examines how cohesion emerges and is sustained in ICE and explores whether it is associated with performance and team composition in this setting.

Taken together, this thesis offers an important theoretical contribution by demonstrating the benefits of considering intra-extreme differences, in addition to the highlighting how ETs differentiate from conventional teams. Distinguishing between types of ET will better direct empirical research to the questions that need answering in each context and will improve the applicability of research as findings in one type of ET (e.g., MTS) can be more readily applied to teams who share similar defining features.

Overview of Thesis Chapters

Due to the related, but distinct nature of emergency response and expedition teams, the decision was made to submit this thesis in the alternative format. This facilitated a finer grained examination of each team and allowed papers to be written in a format suitable for publication. An overall discussion of findings across contexts and a reflection of the methodological/analytical approaches best suited to studying these teams can be found in *Chapter VI*. The first aim of this thesis is to explore what factors support teamwork across two types of ET (MTS and ICE). The second aim of this thesis is to identify the methods and analytical approaches suited to studying two types of ET. This thesis draws on a number of methods in researching ET (*see* Table 1).

Table 1. Overview of Thesis Chapters and methods.

Chapter	Research Focus	Context	Method	Data	Analysis
II	Methods and Analytical approaches	ETs situated in organisational contexts (e.g., critical care, emergency response)	Literature review	N/A	N/A
III	Communication, Coordinating Behaviours	Emergency response teams (MTS)	Simulation	Audio recorded interactions	Social Network Analysis, Thematic Analysis, Chi Square test
IV	Familiarity, Individual differences, Communication, Shared goals, Coordinating behaviours	Emergency response teams (MTS)	Simulation	Audio recorded interactions, self-report survey, decision logs	Social network analysis, Mediation, Hierarchical Regression, Thematic Analysis.
V	Cohesion and Individual differences	Expedition teams (ICE)	Longitudinal diary study	Self-report survey	ANOVA, Multi-level modelling, Bayesian statistics

Chapter II

Chapter II was written for a Special Issue on "The challenges to working with real teams: challenges, needs and opportunities" in Organizational Psychology review. The Chapter serves two key purposes within the over-arching thread of the thesis: (i) it makes the case for conducting further research with ETs, and (ii) it outlines the challenges to existing methodologies before identifying immersive simulations as a useful alternative. Immersive simulations allow the meaningful measurement and analysis of behaviour in an environment that emulates the organisational context of interest (Alison et al., 2013; Manser et al., 2007). With recent methodological, technological and analytical advances in psychological research, this Chapter presents a framework for conducting simulation research with ETs that suggests how to best utilise these advances. Due to the nature of the special issue, no reference is made to expedition teams in this Chapter. Rather, the Chapter focuses on outlining a framework for conducting research with ETs in organisational settings (e.g., emergency response, critical care). This Chapter has been published: Brown, O., Power, N., & Conchie, S. (2020). Immersive Simulations with Extreme Teams. *Organizational Psychology Review.*

Chapter III

Building on suggestions made in *Chapter II*, *Chapter III* is a simulation study that examines teamwork across time in a MTS response to an emergency. This Chapter uses a mixed methods approach to explore how multi-agency response teams communicate and coordinate at three time points following a simulated terrorist incident. Whilst research exists on communication and coordination during emergencies, this work is largely atheoretical with few examples that have studied these processes over time (Chen et al., 2008; Mishra et al., 2011; Salmon et al., 2011). The aim of this Chapter was to address these limitations to prior research, drawing on theoretical frameworks within the MTS literature to make predictions and interpret findings (e.g., Luciano et al., 2018), in addition to exploring how changes in contextual demands drove changes in how teams interacted with one another. Social network analysis was used to identify differences in communication networks over time and thematic analysis of transcribed audio data identified coordinating behaviours and how they changed over time. This Chapter has been submitted to the *Journal of Occupational and Organizational Psychology*.

Chapter IV

Chapter IV explores the effects of familiarity on team processes in MTS. In conventional teams, familiarity is associated with improved performance (Espinosa et al., 2007; Harrison et al., 2003). However, there is a lack of empirical research exploring the role of familiarity in MTS (Turner et al., 2019). This is important as the coming together of previously unacquainted teams is cited as a key barrier to effective teamwork in this context (Shuffler et al., 2015; Waring et al., 2019). An experimental simulation study was carried out to examine the role of familiarity in shaping MTS behaviours. The primary aim of the study was to identify differences in the coordinating behaviours of familiar and unfamiliar teams. The secondary aim was to test if any differences were due to communication structures and shared goals. Social network analysis coupled with thematic analysis of transcripts were used to identify differences in communication and coordinating team behaviours. Mediation analysis was used to test for the indirect effect of familiarity on team behaviours through communication and shared goals. This Chapter has been submitted to *Small Group Research.*

Chapter V

Chapter V is a longitudinal study, piloting a diary methodology to study cohesion over time in expedition teams. This was based on prior research theorising that cohesion will be especially important to ensure effective team performance in isolated, inhospitable environments (Stuster, 2011; Vessey & Landon, 2017). Quantitative diary-based data were collected daily from five teams, travelling to three different locations for a period of 20 days. Linear mixed model analysis was conducted to identify patterns in cohesion over time and to assess the extent to which fluctuations could be explained by variations in daily events (e.g., weather conditions, physical health). Pre-expedition personality measures were used to identify if changes in cohesion over time could be explained by the composition of the teams. This paper has been published: Brown, O., Barrett, E., Power, N. (2019) Monitoring cohesion over time in expedition teams; the role of daily events and team composition. In *Proceedings of the 14th International Naturalistic Decision-Making Conference, San Francisco, U.S.A*.

Chapter VI

The final Chapter of this thesis presents an overall discussion of findings, with reference to the research questions posed in *Chapter I*. As a more focused discussion of each individual paper is provided in the corresponding Chapters, *Chapter VI* focused on emerging themes across Chapters and the limitations, in addition to possible future directions of research into ETs.

Chapter II

Immersive simulations with Extreme Teams

Chapter 1 provided a general overview to the thesis, identifying research questions and reviewing the current literature in the field. This Chapter addresses the second research question of this thesis; *"What methods and analytical approaches are suitable for studying teams in different types of extreme environments"*. This was a focused review article, written for a special issue in Organizational Psychology Review on: "*The Challenges of Working with "Real" Teams; Challenges, Needs and Opportunities"*. As such the ETs referred to within this Chapter are situated within organisations and no mention of expedition teams is included.

As identified in *Chapter 1*, simulations are a useful platform for conducting research with emergency responders and these benefits extend to other ETs that operate within a range of organizational contexts (e.g., military, medical teams). Whilst researchers have had success researching ETs with interviews, observational studies and experimental studies (*see* Alison & Power, 2017a; Gillespie, Gwinner, Chaboyer & Fairweather, 2013; Zaccaro, Gualtieri & Minionis, 1995), simulations have unique benefits to advance understanding of teamwork in challenging and extreme environments. Unlike other approaches, simulations afford a high level of experimental control, in a context that adequately re-creates the challenges of the organisational context (Manser, Dieckmann, Wehner & Rall, 2007). This is especially important with ETs as it allows the meaningful measurement of behaviours in real-time, enabling researchers to test specific aspects of theory in an environment that is both physiologically and psychologically safe (Alison, van den Heuvel, Waring, Power, Long, O'Hara & Crego, 2013).

Simulation research has been carried out for decades (Rosen et al., 2008). With recent technological and analytical advances, it is appropriate to take stock and consider how to capitalise on these advances in simulation studies. For example, discussing how simulation research might make use of the recent advances in Bayesian statistics in psychological research. In the following Chapter, we acknowledge the challenges of researching ETs and review existing methodologies before presenting a framework for conducting immersive simulation research that focuses on; design, data collection and data analysis.

Abstract

Extreme teams work in challenging, high pressured contexts, where poor performance can have severe consequences. These teams must coordinate their skill sets, align their goals, and develop shared awareness; all under stressful conditions. How best to research these teams poses unique challenges as researchers seek to provide applied recommendations whilst conducting rigorous research to test how teamwork models work in practice. In this paper we identify immersive simulations as one solution to this, outlining their advantages over existing methodologies and suggesting how researchers can best make use of recent advances in technology and analytical techniques when designing simulation studies. We conclude that immersive simulations are key to ensuring ecological validity and empirically reliable research with extreme teams.

Keywords: *Teamwork, Simulations, Extreme teams*

Immersive simulations with Extreme Teams

'Extreme teams' (ETs) operate in challenging environments in which there are considerable physical, psychological and interpersonal demands (Manzey & Lorenz, 1998). ETs share many similarities with 'High Reliability Organisations', in which teams are required to operate effectively, in complex task environments, and for sustained periods of time (Roberts, 1990; Klein, Bigley & Roberts, 1995). What both contexts have in common, and what defines an ET, is that they operate in atypical environments (in terms of demands/stress levels), in which ineffective performance can have severe, potentially life or death, consequences (Bell, Fisher, Brown & Mann, 2018). Examples of ETs include those involved in long-duration space flights (Zhang, Olenick, Chang, Kozlowski & Hung, 2018), submarine command and control rooms (Stanton & Roberts, 2018), medical emergencies (Klein, Zeigert, Knight & Xiao, 2006), high-risk industries (Sneddon, Mearns & Flin, 2006) and emergency response (Power & Alison, 2017a). Interest in ETs is increasing (*see* Driskell, Driskell & Salas, 2018; Roma & Bedwell, 2017), with teamwork viewed as a vital component to organisational success and safe working practices (Hughes et al., 2016; Mazzocco et al., 2009). This has led to a consideration of how to study these unique, often hard to reach teams and to conduct rigorous applied research that contributes to wider theoretical understanding (Bell et al., 2018; Kozlowski, 2015). Given the unique context in which ETs operate, this understanding may diverge from what we know about conventional teams and challenge our current thinking. We identify immersive simulations as one way to achieve this and present a framework for designing, conducting and analysing this research, drawing on current research and ethnographic experience.

Researching extreme teams

Teamwork is essential for safety and success in extreme environments (Hughes et al., 2016). For example, research in high-risk industries has shown that accidents occur more often due to problems between team members than unsafe working conditions (Dwyer & Raftery, 1991); a finding that has been attributed to poor leadership (McCabe, Loughlin, Munteanu, Tucker & Lam, 2008) and a lack of team spirit (Kadiri et al., 2014). Risser, Rise, Salisbury, Simon and Berns (1999) also showed from 54 incidents across eight US hospital emergency departments that half of all recorded deaths and permanent disabilities could have been prevented through better teamwork. Identifying solutions to improve teamwork in ETs can be challenging. This is because they have complex team structures, often form (and dismantle) rapidly, draw on multiple agencies and operate in dynamic conditions that impose a high level of stress on members due to the severe consequences of poor teamwork (Crichton, Flin & Rattray, 2000; Schmutz, Lei, Eppich & Manser, 2018). These features are different to what we see in conventional teams and suggest that theoretically, their processes may be structured differently.

Research on teams requires careful consideration of the complex interplay between performance and its antecedent factors that reside at four levels: the individual (e.g., personality), the team (e.g., team structure: horizontal or vertical), cultural (e.g., organisational culture) and contextual (e.g., task demands). Each of these levels, in isolation and in combination, influence how well a team adapts and responds to a situation. When applied to ETs, an extra layer of complexity is added when we consider the extent to which psychological pressures (e.g., stress) interact with each of these levels and alters team performance (Driskell et al., 2018). The experience of stress can create a perception that task demands exceed available resources, which can lead to undesirable physiological, psychological, behavioural and/or social outcomes (Salas, Driskell & Hughes, 1996). These demands may reside in conventional teams to a lesser extent (or not at all), or in a qualitatively different way (e.g., relating to performance rather than the loss of life). Differences in contextual demands can drive the *type* of stress experienced in teams, which may change or amplify the drivers of effective teamwork (Driskell et al., 2018; Maynard, Kennedy & Resick, 2018). Considering this, researchers have called for empirical research with ETs to test if theoretical models developed with conventional work teams apply to those working in these challenging settings (Vessey & Landon, 2017), and to develop solutions that can protect workers and enhance performance (Power, 2018).

Simulation research with extreme teams

Researchers looking at ETs have employed a variety of methods to understand their composition, function and processes. When the research question concerns a descriptive understanding of ETs, qualitative methods such as observations and interviews (used in isolation or together), have been shown to be effective. Gillespie, Gwinner, Chaboyer and Fairweather (2013), for example, developed an ethnographic account of surgical teamwork culture using observations and interviews. Power and Alison (2017a) identified nine core challenges for commanders during emergencies using interviews. When the research question concerns the influence of self-perceptions on teamwork, self-report measures such as questionnaires have been used. Wauben et al. (2011) found differences between medical team members' in the way they perceived non-technical skills (e.g., communication and situation awareness) using a questionnaire survey. However, what these studies do not do, and what is distinct in simulation studies, is manipulate specific variables to test theory and generate empirical evidence of how these variables influence team performance. Whilst the manipulation of variables is possible in traditional laboratory studies, these studies often utilise student samples in a setting that is void of the stressors present in an extreme environment (e.g., Zaccaro et al., 1995). Further, research highlights the importance of expertise in extreme environments (Boulton & Cole, 2016), thus suggesting that understanding how practitioners work in the real-world necessitates that research is undertaken with the population of interest.

One effective method for studying ETs are simulations. Simulations allow for the measurement of complex relationships between factors that impact team performance in a meaningful organisational context, whilst facilitating a high level of experimental control (Alison, van den Heuvel, Waring, Power, Long, O'Hara & Crego, 2013; Manser, Dieckmann, Wehner & Rall, 2007). Example relationships may include the impact on performance of individual differences (e.g., attitudes), trust between team members, temporal patterns in teamwork over time, and cultural and contextual variables that may moderate these relationships, such as organisational norms and task demands. Studies that have used simulations to answer such questions include Bienefeld and Grote (2014) who showed the influence of expertise and organisational knowledge on leadership behaviours in aviation teams; and Amacher et al. (2017) who demonstrated that all-female medical teams showed less "hands-on" time and a greater delay before chest compressions in comparison to all-male teams.

In comparison to alternative methods, simulations have five key benefits; they: (i) re-create the stressors and challenges of the workplace; (ii) involve data collection with the population of interest (i.e., practitioners instead of students); (iii) provide an opportunity for researchers to test theory by manipulating and measuring discrete variables; (iv) allow for the collection of rich quantitative and qualitative data related to team behaviour in real time, and (v) can be used as a training tool to increase participation (Rosen et al., 2008). Simulations are an especially useful platform for collecting data with ETs as they provide a physiologically and psychologically safe space that will not endanger participants (Alison et al., 2013), whilst eliciting similar behavioural patterns as would be found *in situ* (Manser et al., 2007). They are also suited to research with ETs who may be difficult to study using alternative methods (e.g., the security sensitive nature of military command control would negate an observation study).

This paper has two main aims. Firstly, it seeks to show the utility of immersive simulations in studying a range of ETs; not just those who operate in healthcare, where many of the frameworks and benefits of utilising immersive simulations originate (*see* Cheng et al., 2016, Cheng et al., 2014). We will show in this paper that they can also be in contexts where ETs are less well-structured (e.g., multi-team systems), more fluid (e.g., non-stable team members) and involve both horizontal (i.e., within an operational team) and vertical (i.e., between operational, tactical and strategic teams) organisational structures. Secondly, the paper will outline recent technological and analytical advances in psychological research and consider how simulation research can be improved by utilising more immersive methods that can better harness these advances. For example, by considering in what way emerging virtual reality technologies or alternative statistical approaches (i.e., Bayesian statistics) might be used to allow advanced models of ETs to develop. These developments have implications beyond the ET context and hold promise for team research in general. In this paper, we address these aims by outlining a framework for using immersive simulations for research with ETs, broadly focussing on three aspects of the research lifecycle: (i) simulation design, (ii) data collection; and (iii) data analyses.

Simulation design

A simulation seeks to create a testing environment that closely replicates reality (Sleeper & Thomspon, 2008). An important consideration during research design is how to embed fidelity and immersion so that participants feel engaged in the simulation and exhibit similar behaviours as would be found *in situ*. Fidelity and immersion are two inter-related constructs that seek to increase the sense of realism during a simulation (Alison et al., 2013; Lester, Georgiou, Hein, Littlepage, Moffet III & Craig, 2017), and which determine the success of simulations. *Fidelity* is the extent to which the simulation matches the real-world environment (Maran & Glavin 2003). This influences the level of *immersion* felt by the participant, defined as the "subjective impression that one is participating in a comprehensive, realistic experience" (Dede, 2009, p. 66). Fidelity can be created at the physical and psychological levels. Physical fidelity refers to the extent to which the simulation reflects the material aspects (i.e., a physical replica) of the working environment (Lester et al., 2017). It is based on the principle that the more similar the simulated task environment is to the real environment, the greater the transfer of learning (Baldwin & Ford, 1998). Psychological fidelity refers to the degree to which the skills and behaviours necessary to complete organisational tasks are accurately represented in the simulated environment (e.g., does the task evoke a similar level of cognitive processing) (Bradley, 2006). Psychological fidelity is expected to elicit similar psychological processes necessary for real-world performance (Kozlowski & DeShon, 2004). The decision on whether to maximise physical fidelity, psychological fidelity, or both during research design is dependent upon the research questions of interest.

Physical fidelity is important when a level of "dexterity" is needed by the target population to complete the task (Dieckmann, Gab & Rall, 2007). It allows the transfer of procedural skills that might not be possible using psychological fidelity methods alone (Hochmitz & Yuviler-Gavish, 2011), and is especially important when the research question concerns an interplay between humans and hardware (e.g., does a new piece of kit promote faster teamwork?). Understanding the interplay between humans and hardware, referred to as a "sociotechnical system" (Baxter & Sommerville, 2011), is important for ETs as their context becomes increasingly digitised. ETs where this will be important include control room operators, flight-crews, and emergency medical teams. For example, Stachowski, Kaplan and Waller (2009) used an exact replica of a nuclear control room to study adaptability of teams as they moved through the testing space, communicating and sharing information with colleagues whilst interacting with the electronic displays to rapidly find faults and implement changes to systems (Waller & Kaplan, 2018). Although essential for certain ETs (e.g., operational teams that need to interact with hardware), creating physical fidelity through physical replicas can be difficult as they are often expensive, take up a large amount of physical space, and are often not portable (Kozlowski & De Shon, 2004).

Psychological fidelity is important for researchers interested in studying non-technical skills in ETs (e.g., trust, decision-making, sensemaking), or teams operating at strategic levels. It allows for the examination of the interplay between individual and contextual factors on intra-team processes (Kozlowski & DeShon, 2004). For example, researchers interested in the effects of psychological stressors (e.g., task-related anxiety) on team communication and coordination might build reactionary consequences into the simulation design to increase the gravity of decisions and sense of accountability of decision-makers (Eyre, Crego & Alison, 2008). This might be achieved by gathering team members round a board room style table and providing them with real-time information that follows a realistic narrative to an unfolding situation (e.g., video calls from simulated team members, PDFs with 'data' related to the simulation exercise). An example of where this has been used successfully is Power and Alison (2017b). They ran a simulation study examining how a team of emergency service commanders made decisions during a simulated terrorist incident in which different injects were presented to team members dependent on their answer during the previous inject. This enabled participants to feel immersed by embedding consequences for choices, increasing the gravity of decision-making.

Recent advancements in virtual reality (VR) software offers an accessible and highly immersive way to achieve both physical and psychological fidelity. VR are "computer-generated simulations of three- dimensional objects or environments with seemingly real, direct or physical user interaction" (Dionisio & Gilbert, 2013, p2). They offer an affordable alternative to physical replicas of the organisational environment, whilst still testing important teamwork processes in a context that mirrors the decisions and challenges present in the workplace (Pan & Hamilton, 2018). VR simulations can therefore be used to test both operational (e.g., physical tasks) and strategic teamwork (e.g., decision-making).

One example of a VR system is the Cave Automated Virtual Environment (CAVE). CAVE comprises an enclosed cube, sitting within a large darkened room with projectors on each side (Cruz-Nierra, Sandi, DeFanti, Kenyon, & Hart, 1992). CAVE is attractive as the goggles that are worn do not stop participants from seeing their own hands (as with most head mounted VR devices), whilst they interact with the VR projected on the screens. This means that participants can interact with physical objects (e.g., enact driving by using a real steering wheel) (Pan & Hamilton, 2018), allowing researchers to examine the ability of teams to perform physical tasks. This is especially important when researching ETs that are required to complete arduous physical tasks (e.g., search and rescue teams), and may offer some insight into how contextual demands can influence team members' ability to use specialist equipment. For example, CAVE has been used to train firefighters using Breathing Apparatus Entry search methods - searching a building for casualties in which sight and breathing is restricted by smoke (Backlund, Engstrom, Hammar, Johannesson & Lebram, 2007). In their study, participants wore personal protective equipment and sensors were fixed to the walls so that physical movements within the "CAVE" corresponded to their movement and orientation within the simulation. This increased the physical effort needed to complete the tasks, giving participants a sense of real-world orientation whilst in a virtual world.

The use of CAVE is not widespread generally (*see* Jiang, Rahimian, Yon, Plumert, & Kearney, 2016) and this is especially so in relation to ETs. This may be attributed to the fact that it is relatively expensive and difficult to transport in comparison to other VR systems such as head mounted displays (Mallaro, Rahimian, O'Neal, Plumert, & Kearney, 2017). However, evidence from other areas have shown its potential utility for understanding ETs. Gamble et al. (2018) utilised the CAVE system to explore friend/foe discriminatory fire in military personnel, where they found that participants made more errors when under stress, but that 'expertise' was a protective factor. There is also evidence from its use in social psychology that it may be used to explore the role of social influence on individual behaviour. For example, Kinateder et al. (2014) showed that the presence of a virtual agent significantly affected route choice in the evacuation of a tunnel fire. Applied to ETs, the potential for unpacking social influence suggests that CAVE may help develop our understanding in areas such as how intra- and inter- team communication influences performance in multi-team systems (MTS). At present there is limited understanding of how behaviours at the intra-team level affect inter-team performance and vice versa (Asencio & DeChurch, 2017). In the immediate term, and commensurate with the current capability of this kit, we would expect ET research utilising this technology to focus on questions that do not require data to be collected from multiple team members in parallel. In the longer-term, and as this technology advances, we see potential for the CAVE system to study the interaction between multiple individuals, in addition to its current capability of studying the interaction between participants and virtual agents.

When designing a simulation, the involvement of practitioners and/or experts is invaluable. They can ensure that the simulation is relevant to organisational tasks (Klein & Woods, 1993), provide expert input about the task environment and narrative, increase the likelihood that the simulation will elicit similar cognitive and emotional responses found in the real world (Crandall, Klein, Klein & Hoffman, 2006), and help to ensure that simulations offer *both* research and training benefits, which can facilitate participant recruitment and engagement (Rudolph, Simon & Raemer, 2007; Waller & Kaplan, 2018). This makes simulations attractive to end-users as they provide a space in which team skills can be trained, facilitating recruitment that might otherwise be challenged by the high workloads and small populations of participants (Beaubien & Baker, 2004).

However, it is important to ensure a balance is met between research and training goals. Simulations can be resource intensive and it is important that researchers are not prevented from collecting the data they need to answer their research questions and practitioners are not promised a simulation that fails to meet their training objectives. To do this, researchers must delineate what the training goals of the organisation are during the early phases of design, and work around them to ensure training objectives are compatible with research goals (Dieckmann et al., 2007). This should facilitate an interdisciplinary partnership and enable collaboration through the entire research project. The involvement of practitioners at the early stages of research can also have benefits later on in terms of research dissemination and impact. Practitioners are keen to receive feedback on their training, as such, a research team might want to organise a feedback workshop or write a practitioner-friendly report on findings. This can facilitate opportunities for further follow-up studies and ensure a collaborative relationship with practitioners moving forward.

Data collection

A key benefit to simulation research is that it facilitates the collection of rich behavioural data, allowing researchers to study the verbal and non-verbal dynamics of teamwork. Psychology has seen a decline in the use of behavioural measures in recent years, typically showing a tendency to use self-report surveys (Cialdini, 2009; Dolinski, 2018). However, there has been a general call to move beyond self-report measures to gain a better understanding of how social coordination emerges in complex environments (Willemsen-Dunlap et al., 2018) and to develop more objective measures of behaviour (Rosen & Dietz 2017). This is due, in part, to the limitations of solely using self-report measures which; (i) fail to account for the richness of team-based interactions (Shuffler & Carter, 2018); (ii) lead to a proliferation of scales each attempting to measure the same thing (*see* Salas et al., 2015 on team cohesion); (iii) show weak correspondence with non self-report outcome measures (*see* Valentine, Nembhard & Edmonson, 2015 for a review in a health care setting), and (iv) are subject to a number of biases (e.g., Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). We suggest that simulations offer a methodological advantage to self-report by recording behaviour *in situ.*

Wearable technology. The tools used to collect data during simulations need to be unobtrusive so as not to break immersion, but robust enough to allow for reliable examination of the research question. The advancement of behavioural measures creates promise for the use of wearable sensors in research using simulations. Wearable sensors are mobile devices that record data on how the wearer interacts with their surroundings (including other people). They do this using microphones, accelerometers, infrared sensors and/or Bluetooth components (Chaffin et al., 2017). Wearable sensors have advantages over traditional methods; namely that they allow for the effortless recording of data from participants that are not reliant on self-reports, and that data are real-time and continuously collected thus removing the necessity for researchers to piece together static data taken at set times, sometimes from multiple devices. This makes wearable sensors especially suited to simulations, as the continuous collection of rich data in the real world may lead to consent and confidentiality issues (e.g., recording patient-clinician interactions).

The fact that behavioural data are collected continuously means that wearable devices have the potential to identify important *within*-person insights and their impact on team performance. This has not always been achieved with traditional methods, which tend to focus at the *between*-person level (Matusik, Heidl, Hollenbeck, Yu, Lee & Howe, 2018). This finer grained understanding of how teams operate has the potential for simulation methods to develop complex, non-linear, relationships between relational variables. For example, data from wearable sensors may allow for the development of a finer-grained understanding of leadership in ETs, such as how a leader's behaviour fluctuates across an emergency and how these fluctuations impact behaviours. Similarly, it may examine how leadership changes interact with team factors (e.g., the presence of other teams – as within MTS) or external forces (e.g., contextual demands – during crises response).

At a theoretical level, wearable sensors are most valuable when the research question concerns relational issues at the team level (e.g., cohesion, trust, leadership), as they show how the person navigates their environment, including social interactions. In using data from single or multiple streams (e.g, audio, Bluetooth), studies have used wearable technology to examine affect and team cohesion in simulated space exploration missions (Zhang et al., 2018), cooperation (Taylor, 2013), communication in productive and creative teams (Pentland, 2012), social and task-related exchanges (Matusik et al., 2018), social networks (Wu, Waber, Aral, Brynjolfsson & Pentland, 2008), boundary spanning individuals (i.e., those that coordinate activity between established groups) and emergent leaders (Chaffin et al., 2017). There is potential for research in ETs to build on this to use sensors in the study of MTS, to explore how boundary spanning individuals support teamworking across multiple agencies responding in crises. Previous research has tended to rely on self-report and coding of verbal behaviours (*see* Bienefeld & Grote, 2014), whereas wearables can measure other aspects such as variations in proximity over time (i.e., using Bluetooth), in addition to providing a continuous measure of communication.

Research using audio data more generally expands the potential for wearables in simulation research. For example, Stanton and Roberts (2018) used audio data to understand team level macrocognition (i.e., cognitive functions that are performed in naturalistic settings, *see* Klein, Ross, Moon, Klein, Hoffman & Hollnagel, 2003), Bowers, Jentsch, Salas and Braun (1998) have used it to understand shared mental models, and Fischer, McDonnel and Orasanu (2007) have used it to examine which *types* of information (task or relational focused) best support performance in ETs. From the perspective of understanding ETs, this is especially promising as the nature of these environments means that team members have to share, analyse and discuss complex information (e.g., Haddow & Bullock, 2003). An important question for ETs, due to the time sensitive nature of their work, is how to do this efficiently. Evidence from a range of non-ETs suggests that short and equal verbal contributions, face-to-face communication, distributed connections within the team and information seeking from other teams characterise success (Pentland, 2012). Wearable sensors would allow for a reliable test of these hypothesised effects in ETs, whilst maintaining the realism of the ET environment through the use of the simulation.

Wearable technology can be used to record physiological data from team members. Psychological pressures (e.g., stress) is an important factor to consider in simulation research of ETs as the inherently stressful environments they operate in can disrupt performance (Driskell et al., 2018). Stress has been shown to reduce communication in aviation teams (Sexton & Helmreich, 2000), impair cognitive functioning in military teams (Wallenius, Larrson & Johannsson, 2004) and reduce information sharing in less experienced surgical teams (Wetzel et al., 2006). Although previous research has explored the role of stress in ETs, studies have often failed to check whether the experimental manipulation has actually affected stress levels or, alternatively, used a self-report survey to do so. For example, increasing stress by imposing time pressure has been associated with an increase in risk-taking behaviour (Young, Goodie, Hall & Wu, 2012), and a shift towards more satisficing decision styles (Alison, Doran, Long, Power & Humphrey, 2013). However, neither of these studies took physiological measures of stress from their participants and so the effects of stress, via time pressure, were assumed.

Wearable technology allows us to address the limitation of these other studies. It is possible to measure stress during a simulation by using wearables that record 'stress-related' measures, such as heart rate, galvanic skin responses and change in pitch (Mozos et al., 2017). For example, stress during a simulated driving task, as measured using skin conductivity (i.e., sweating) and heart rate, has been found to predict stress levels with the highest level of accuracy when compared against physical indicators of stress (e.g., breaking and sharp turning) and self-report measures (Healey & Picard, 2005). Heart rate has also been identified as the best indicator of stress in a study comparing physiological indicators of stress during a simulated virtual environment that invoked fear by placing participants over a chasm at great height (Meehan, Insko, Whitton & Brooks Jr, 2002). When applied to ETs, physiological indicators of stress open the possibility of building models that map team responses across a stress episode: from its origin through peak to end. What sets these models apart from conventional teams (where such devices are equally insightful) is the potential for ET models to overlay the stress episodes experienced by inter-related teams (e.g., MTS) to examine interplay or contagion.

In keeping with the need to maintain fidelity during simulations, researchers may also consider using physiological measures of stress to provide an objective indication of how immersive a simulation has been. Baker et al. (2017) used a heart rate monitor to assess if the stress experienced in medical procedures could be replicated within a simulated environment and found that the simulated procedure did not accurately re-create the same level of stress as experienced within hospitals. This emphasises the need to incorporate a physiological measure of stress to ensure that elements of the simulation that are intended to be difficult induce a level of urgency within the participants. There is currently a lack of research that has sought to establish what stress levels are needed to ensure that simulations are useful for training and research purposes (Cumin, Boyd, Webster & Weller 2013). More research is needed to establish standardized levels of immersion which can leave organisations confident that simulations are achieving their intended purposes (Cumin, Weller, Hender & Merry, 2010).

Interactions within the simulation system. Although wearable sensors have the potential to provide rich data on relational issues at the team level, they may not be able to provide a holistic overview of teamwork, such as when communication occurs via other mediums (e.g., email) or when inter-dependent tasks are carried out in different locations. For example, some ETs (e.g., MTS) will operate across several sites and researchers may wish to explore how cultural factors (e.g., organisational policies) and team structures facilitate/hinder inter-team processes. One benefit of simulations is that teams are operating in designated room(s), and so forms of data collection can be built into the simulation system to provide a comprehensive account of verbal and non-verbal communication between team members. Data gathered from participant interactions within the simulation system might include video recording, for example, CCTV of the team operating in the simulation room; or recording data within the simulation computer system itself, for example, by generating a log of clicks or button pushes when participants interact with the simulation; collecting time-stamped 'decision logs'; and eye tracking on the computer screen. Monitoring the interaction within the simulation system may prove particularly important for researchers interested in designing a simulation with high physical fidelity to explore sociotechnical systems (e.g., how team members interact with the computer system). Future research could consider how simulations with high physical fidelity might advance theory on sociotechnical systems and their use by ETs. For example, in considering the role of the team in increasingly automated systems or in what way do contextual demands (e.g., dynamic task requirements) impact team members' ability to effectively utilise technology in crises.

The type of data recorded in the simulation system will be dependent on the system being used and the research questions of interest. For example, research questions that are interested in how team-level factors (e.g., composition) influence decision speed might use a time-stamped 'decision log'. Power and Alison (2017b) used this method to identify how long it took teams to make decisions and how this interacted with the team's goal. Teams were requested to 'log' their decisions on a computer when they wanted to make a decision and these data were automatically recorded and timestamped in the simulation system. Alternatively, researchers may use the simulation system to monitor how team members communicate electronically with one another. Alison et al. (2015b) were interested in communication patterns between sub-teams in different 'syndicate' rooms in a simulation. To do this, they built a 'chatbox' function into the simulation system so that sub-teams could communicate between rooms, with all electronic communications data recorded and time stamped. The simulation system therefore offers an alternative mode of data collection that can be used in isolation or in conjunction with wearable devices dependent on the research question.

Data Analysis

Simulation research with ETs has the potential to yield vast amounts of data from multiple sources, measuring multiple variables. It is important that data analysis maximises understanding of this rich data. There exists a number of methods of analysis that can be used. Here, we focus on two types that are especially relevant: (i) network analyses, which examine interpersonal dynamics within a team at a single time point; and (ii) temporal analyses, which track interpersonal dynamics over time. We focus on these methods as they provide rich representations of *team interactions*, as oppose to assessing the *individual* performance of team-based skills (e.g., Yule et al., 2008). We then turn our attention to the possibility of using Bayesian statistics, which allow analyses to be carried out with smaller samples, and thus may open up the possibility of testing more complex theoretical models in ET research.

Network analyses. Network analyses allow a researcher to analyse team behaviour during simulations by quantifying information and providing a visual representation of how team members interact. This type of analysis is especially useful when comparing how contextual factors (e.g., task type) influence team behaviours (e.g., inter-team communication) (Stanton & Roberts, 2018). Using recorded communication data (e.g., by using wearable devices or CCTV recordings), Social Network analysis (SNA) shows how team members communicate with each other and the centrality of any one member (Knoke & Yang, 2008). SNA are also useful as they provide a visual representation of the social dynamics of a team by plotting each person as a node and showing the strength of the connections between them. At a theoretical level, this is especially important for ETs that involve multiple agencies operating within a hierarchical structure as it can identify instances in which communication patterns do not follow pre-defined organisational processes and structures (Dekker, 2000), or plausible reasons for communication breakdowns. For example, SNA has been used to identify key tasks that challenged communication in submariners (Stanton & Roberts, 2018), how team communications varied dependent on team composition in surgical operating staff (Anderson & Talsma, 2011), and how a lack of connectedness between a Search and Rescue Team contributed to faulty communications and the ability to develop shared situation awareness (Fodor & Flestea, 2016).

An alternative type of network analysis that goes beyond communications data is the Event Analysis of Systematic Teamwork (EAST) technique. This method models the macrocognition (i.e., situation awareness) of a team by generating task and information networks in addition to social networks (Walker, Gibson, Stanton, Baber, Salmon & Green, 2006). In order to perform EAST, raw data from audio and video recordings are transcribed and then used to create matrices of each of the three networks (i.e., social, task, information). This results in a "network of networks", that allows researchers to identify how constructs in different networks might interrelate. For example, communications might influence the way a task is performed, which might influence how information is transferred.

EAST has been used to examine teamwork in simulation research across several extreme contexts; submariner command and control (Stanton & Roberts, 2018); emergency response (Houghton, Baber, McMaster, Salmon, Stewart & Walker, 2006) and air traffic control (Walker et al., 2006). As EAST involves generating a task network, it is useful for researchers who are interested in understanding how team members coordinate to complete tasks as well as how they communicate with one another in extreme environments. Hierarchical Task Analysis is a methodology within EAST that is used to identify key tasks (Annett & Stanton, 2000), as well as the individuals who complete tasks, the structure, and the order in which the tasks take place (Walker et al., 2006). This provides a detailed representation of how team goals interact and are resolved (Walker, Stanton, Baber, Wells, Gibson, Salmon, & Jenkins, 2010). For example, a simulation researcher interested in team coordination may want to model how a team approaches different tasks dependent on difficulty. As coordination is defined as the behavioural mechanism enabling teams to sequence, synchronize and integrate their efforts in order to achieve goal-relevant tasks (Marks, Mathieu & Zaccaro, 2001), modelling how teams move through tasks should contribute to a more complex understanding of how ETs coordinate. This is extremely relevant for researchers interested in ETs due to the importance of coordination in manging complex team structures and preventing error across a range of contexts such as aviation (Grote, Kolbe, Zala-Mezo, Bienefeld-Seall & Kunzle, 2010) and medical emergency teams (Schmutz & Manser, 2013).

Temporal analysis. Temporal analysis seeks to identify how team behaviour might change over time in response to changes in individual, team and contextual demands. This type of analysis is especially useful for ET researchers interested in exploring how team processes emerge and are sustained during simulated tasks. It recognises the important role of context in shaping team-based interactions (Ilgen, 1999), emphasising that teamwork does not exist in a vacuum and team processes will change over time (Kozlowski & Ilgen, 2006). Non-simulation based team research has sought to study how teamwork changes over time by collecting longitudinal data (e.g., questionnaires) at set intervals over a given period (*see*, Mathieu et al., 2015). However, this staged approach might not be feasible for some ETs as team members rotate and might not work together at set regular intervals (e.g., emergency response teams). Moreover, these approaches tend to rely on self-report data, as oppose to monitoring actual behaviour in real-time, which has limitations as detailed above (Shuffler & Carter, 2018).

An alternative approach is to study how team behaviour evolves during a simulation. Although simulations will not produce 'longitudinal' temporal data in the traditional sense (e.g., over a course of weeks/months/years), simulations offer a closer replica of how ETs operate in the real world, wherein they must adapt and evolve their teamwork during a given task (e.g., emergency incident). As such, simulations allow us to study the temporal dynamics of teamwork during a simulated "event", which can incorporate multiple goal directed tasks and episodes (Marks et al., 2001). By analysing simulation data longitudinally (i.e. over the course of the simulation), researchers can explore how teams adapt and change as they cycle through different episodes within the simulated event (Marks et al., 2001). The advent of wearables and advancements here allows for this to be done in a reliable and highly detailed way, enabling researchers to begin examining complex, non-static theories or models of behaviour. This could be especially important to advance understanding of MTS. For example, wearable devices may be used to measure communication and relational emerging variables such as cohesion across multiple component teams. When coupled with repeated SNA this would allow researchers to map how intra- and inter- team behaviours and relationships change over time. This could answer questions such as how intra-team behaviours relate to inter-team performance or how intra-team cohesion affects how inter-team members relate to one another.

Beyond comparing networks analyses during different phases of a simulation, a more complex way of analysing temporal data is by using lag sequential analyses, which seek to identify non-random patterns of behaviour during a task (Becker-Beck, 2001). It is useful for research questions that seek to identify how specific team behaviours (e.g., shifts in communication patterns across team members) develop and change over time (Leenders, Contractor, & DeChurch, 2017) and how specific patterns of behaviour can lead to better team performance (Kauffeld & Meyers, 2009). An example of how lag sequential analyses have been used to study ETs during simulations is Cohen-Hatton, Butler and Honey (2015). Their research sought to identify whether commanders in the Fire Service prescribed to the standard decision model used by the Fire Service, or whether they deviated. Participants were asked to "think aloud" (i.e., verbalise their thoughts) during a simulation, and transcripts were coded to identify if participants progressed through the prescribed model of "situation assessment" to "plan formulation" to "plan execution". They found, using lag-sequential analyses, that participants did not follow this pattern. However, a simple goal-oriented training intervention made participants more likely to adopt the prescribed processing pattern, without delaying decision speed. Lag sequential analyses are thus useful for helping to understand patterns in team processing and behaviour during a simulated event, and also provides possibilities for testing interventions to increase adherence to decision models and/or improve performance. For example, using this technique, research might develop our understanding of how patterns of behaviours change dependent on information flow, level of stress in team members (as measured using physiological markers), changes in goal hierarchies; and the interaction between these variables. In doing so, we would have an enhanced understanding of the temporal and contextual influences on teamwork in ETs.

Bayesian statistics. Another approach to analysing data from ET simulations is by using Bayesian statistics. Unlike network and temporal analyses, Bayesian statistics are not a type of data analysis, but are an alternative statistical approach to classic significance testing. Traditional research on teams often draws on classic significance testing (e.g., null hypothesis testing, p-values, confidence intervals) to test specific variables and theories. However, this approach is problematic when working with ETs as, at a practical level, it often calls for moderate to large sample sizes with normal distributions (*see* Wagenmakers et al., 2018 for other problems with classic theory). Research with ETs tends to involve small sample sizes as the participant pool is much smaller than the general population and participants often have limited time to take part in research (Bell et al., 2018). Whilst efforts to address this have drawn on using trainees from ETs, such as trainee paramedics (e.g., Amacher et al., 2017), these samples have been shown to operate differently to "experts" (Boulton & Cole, 2016). In other types of ETs (e.g., emergency response, command and control), trainees may also not be as readily available as they are in clinical settings.

In response to problems with classic testing, researchers are calling for alternative methods of analysis (e.g., Vandekerckhove, Rouder, & Kruschke, 2018). One that has seen an increase in popularity—facilitated by advancements in computer algorithms and quicker hardware processing—are Bayesian Statistics (for example, *see* Special Issues in *Journal of Mathematical Psychology* 2016, vol. 72; *Psychonomic Bulletin & Review* 2018; vol. 25). As a set of tools, Bayesian statistics are attractive to ET simulation research as they open the potential, *inter alia*, for theoretical models to be tested even when samples are smaller than with conventional team research.

As a very broad (and somewhat simplified) overview, Bayesian statistics have the ultimate goal of showing the probability that the data observed is likely to occur under two competing theoretical (i.e., statistical) models (Kruschke & Liddell, 2018). Using Bayes factors, a researcher infers the level of support for their theory, relative to the alternative theory, based on how much the observed data differs from that predicted. This is done by comparing the statistical model against a 'posterior' probability distribution, which is made up from prior information known before data were collected and what is known from the actual – observed – data. Prior knowledge can come from theoretical frameworks, findings from previous research, subject experts and pilot work (Zyphur & Oswald, 2015). Research may also use non-informative priors where knowledge is limited and parameters are set to cover a broad range of possible outcomes, but this is less advisable when samples are small (*see*, McNeish, 2016). Bayesian statistics regard parameters (e.g., probabilities) as variables, and as such, parameters are adjusted as data accumulates and output is compared against starting values. The researcher can thus see how evidence for their theoretical (statistical) model changes with new data; something that is not possible with classic theory where parameters are regarded as constant (*see*, Gelman et al., 2014 for a statistical overview of Bayes analysis; Lynch, 2010 for a general introduction; Jeffreys, 1961 for original writings).

Classical significance tests require researchers to specify in advance what the smallest effect size of interest is given their theory, in order to recruit a sufficient number of participants capable of detecting such an effect. Yet, it has been shown using Bayesian analyses that a high powered non-significant result might not necessarily constitute evidence for the null, and that low-powered non-significant results are not necessarily insensitive (Dienes, & McLatchie, 2018). Evidence suggests that sample sizes estimated using parameters generated through Bayesian analysis rather than power, may be more flexible and yield smaller sample size requirements (Sambucini, 2017). Relatedly, Bayesian analysis has the benefit of allowing for 'optimal stopping'. In essence, this allows for a researcher to track results as data are collected and stop data collection when a certain level of evidence showing one theory as more favourable has been obtained (Kelley, 2013). In addition to allowing for potentially smaller samples to be tested to obtain an effect, this also avoids the ethical issue of testing ET members beyond what is needed.

Bayesian analyses have been applied to a number of methods from t-tests through to structural equation modelling (Brown, Barrett & Power, 2019; McNeish, 2016). For ETs, it could be applied to existing methods (e.g., using a t-test to compare two sets of SNA across phases of a simulation) to identify significant effects that may have been masked by small sample sizes. Depending on the complexity of the theoretical model, we may start to move towards unpacking the different pathways through which factors have an effect on team performance, and the conditions that moderate these effects. This could be especially important in understanding the complex inter-play between component team and system level variables in MTS, which linear approaches may not be able to account for (Cronin, 2015). As interest in larger multi-agency teams expands, we may see the use of Bayesian methods grow as researchers seek to test theoretical frameworks that span multiple levels (i.e., variables at the component team and system level) that traditional statistical approaches would not have the power to do when working with small sample sizes (Wang & Hanges, 2011).

Conclusion

Teamwork is a necessity in almost any twenty-first century organisation, with teams increasingly viewed as the solution to solving complex problems (Salas, Shuffler, Thayer, Bedwell & Lazarra, 2015). This is especially so in organisations operating in extreme environments, where team members must coordinate their behaviour effectively in order to avoid the severe, often life or death, consequences of poor performance. In this paper we have identified the benefits to conducting simulation research with ETs, showing how they differ from existing methods. Second, we have presented a framework for conducting immersive simulations, focusing on three broad aspects of; (i) study design, (ii) data collection and (iii) data analysis. By doing this we have reviewed existing simulation research, as well as suggesting how emerging technologies (e.g., wearable devices, CAVE) and statistical methods (e.g., Bayesian) might be used in simulation research to advance understanding. It is hoped that this paper will inspire researchers to make use of novel immersive simulation-based methods to engender the much-needed empirical research on ETs.

Chapter III

Communication and Coordination across time in the multi-team system response to a simulated terrorist incident

Chapter II outlined the benefits of immersive simulations in the study of ETs. Whilst simulations have been used widely in the study teams (*see* Marks, 2000), *Chapter II* shows how researchers can harness the latest developments in technology and analytical techniques (e.g., Virtual Augment reality systems, Pan & Hamilton, 2018; Social Network analysis, Alison, Power, van den Heuvel & Waring, 2015) to improve the contribution of simulation studies and generate a more complex understanding of ETs. *Chapter III* reports on an immersive simulation study with strategic commanders in the emergency services. It addresses both questions in this thesis. First, the study examines factors that support teamwork in extreme environments by exploring how communication and coordination change across time in a multi-team system (MTS) response to an emergency. It achieves this through qualitative and quantitative analysis of 195 minutes of audio data collected from three simulations with 30 strategic responders. Each simulation emulated a different time point following a terrorist incident (incident ongoing, 48 hours after, 3 weeks after). Second, the study illustrates the points made in *Chapter II* around the use of immersive simulations in the study of MTS. Specifically, this Chapter shows the benefits of: (i) involving practitioners in the simulation design; (ii) using an expert sample; (iii) collecting behavioural data, and (iv) using network analyses across time to evaluate team behaviours.

This study extends previous research in two ways. Firstly, although the MTS literature has advanced in recent years with the development of theoretical models and frameworks, there remains little empirical research that outlines how these models work in practice (Luciano, Mathieu & DeChurch, 2018; Shuffler, Jiménez-Rodríguez & Kramer, 2015). Much of the research in the context of emergency response teams has been atheoretical and applied in focus (e.g., Chen, Sharman, Chakvarati, Rao & Upadhyaya, 2008; Salmon, Stanton, Jenkins & Walker, 2011). By approaching the study of emergency response teams from a MTS perspective, this Chapter offers empirical validation of how communication and coordination unfolds in this context. Taking a MTS approach also ensures that findings have applicability beyond emergency response teams and may inform teamworking in other multi-agency teams (e.g., medical emergency teams, disaster response).

The second way that this Chapter extends prior research is by exploring the effect of time on communication and coordination in MTS. There have been continued calls in the MTS literature to recognise temporal influences on team-based interactions (Luciano et al., 2018; Shuffler et al., 2015). This is because MTS are highly fluid structures, formed in dynamic and complex environments where behaviours are expected to continually change dependent on contextual demands (Aiken & Hanges, 2012; Luciano et al., 2018). Despite this, most MTS research has studied team processes at a single point in time, meaning we have a limited understanding of how processes change and under what conditions different behaviours are optimal (e.g., Bienefeld & Grote, 2014; Luciano et al., 2018; Salmon et al., 2011; Waring, Alison, Shortland, & Humann, 2019). This Chapter addresses this by examining communication and coordination at three time points. It allows a more detailed understanding of how these processes manifest at different points in an emergency and how contextual challenges drive changes in how MTS interact. Identifying which periods of the response are most vulnerable to breakdowns in communication and coordination is important to develop targeted training interventions to improve teamworking in this context.

Abstract

Knowledge of how Multi-Team Systems (MTS) operate outside of the laboratory and across time is limited. To address this, we explored the effect of time on communication and coordination in a MTS response to an emergency. Data were collected from three Strategic Coordinating Groups (SCGs; *N*=30 responders) from 11 agencies who were tasked with responding to a simulated terrorist incident. We collected audio data at three time points that were simulated into the exercise: incident ongoing, 48 hours following-, and three weeks following- an emergency incident. Social network analyses showed that communication networks became less centralised over time, reflecting the fact that once the immediate threat of the incident had subsided, more agencies were involved in decision-making. Thematic analysis of communications between participants identified three positive coordinating behaviours: *joint decision-making, sharing resources* and *sharing task related information*, and three negative behaviours: *role uncertainties*, *decision uncertainties* and *conflicting priorities*. Results showed that positive behaviours increased over time and negative behaviours decreased. Notably, findings suggest the increase in positive team behaviours coincided with increased involvement of additional agencies. Practitioners should consider implementing decentralised networks during emergencies to increase opportunities to share information and to coordinate decisions across inter-agency partners.

Keywords: *Multi-team Systems, Communication, Coordination, Extreme Teams*

Communication and coordination across time in the MTS response to a simulated terrorist incident

Recent events have emphasised the importance of a timely and effective response to large scale emergencies. Whether it be localised incidents such as the Manchester Arena terrorist incident, national incidents such as the Australia bushfires, or global incidents such as the spread of COVID-19, public safety is reliant on effective multi-agency response. Such responses are made by multi-team systems (MTS), which are networks of component teams working to achieve separate but related objectives within a framework of collective over-arching shared goals (Mathieu, Marks & Zaccaro, 2001; Shuffler, Jimenez-Rodriguez & Kramer, 2015). MTS are marked by diverse skills and abilities, and as such, are suited to responding to dynamic, challenging, and complex emergency incidents (Marks, DeChurch, Mathieu, Panzer & Alonso, 2005; Zaccaro, Marks & DeChurch, 2012). However, the fact that MTS comprise different agencies, often with their own priorities and organisational norms, creates sources of error relating to communication and coordination (Fodor & Flestea, 2016; Kerslake, 2018; Pollock, 2017; Salmon, Stanton, Jenkins & Walker, 2011; Waring, Alison, Shortland, & Humann, 2019). Initiatives within the U.K and beyond have sought to improve inter-agency working through the introduction of joint-working principles and increased training, yet problems continue to be identified (JESIP, 2013; Kapacu, 2006; Kerslake, 2018; Waring et al., 2019).

Research focusing on MTS has taken great strides in developing theoretical models that account for the unique challenges facing multi-agency teams, and the potential sources of error (*see* Luciano, DeChurch & Mathieu, 2018; Mathieu et al., 2001; Shuffler & Carter 2018; Zaccaro et al., 2012). Despite their promise, many of these theories are yet to be applied at an empirical level, where research remains largely atheoretical (Chen, Sharman, Chakvarati, Rao & Upadhyaya, 2008; Salmon et al., 2011). There is also a limited consideration of time within much of this applied work, which is surprising given that MTS form quickly, change membership at different stages of response, and continually adjust their behaviours dependent on evolving contextual demands and system objectives (Aiken & Hanges, 2012; Bienefeld & Grote, 2014; DeChurch & Marks, 2006; Luciano et al., 2018; Shuffler et al., 2015). As such, what may mark effective decision-making at one time point may mark an ineffective response at a different time point. Indeed, teamwork research more generally has emphasised the importance of studying teams as dynamic systems that emerge and adapt over time during different performance episodes (*see* Kozlowski, 2018; Kozlowski & Klein, 2000). The current Study starts to address these two limitations by exploring MTS communication and coordination outside the laboratory, and at three different time points during a response to a simulated terrorist incident. In doing this, this Study contributes to the theoretical literature on MTS, in addition to highlighting practical solutions to improve teamwork in this context.

MTS as complex structures

MTS are complex structures that require effective collaboration both within and across teams (Rico, Hinsz, Davison & Salas, 2018; Zaccaro et al., 2012). Compared to conventional work teams, MTS have additional structural features to manage relating to composition (e.g., how many/what kind of component teams are in the MTS), linkage (e.g., how the component teams relate to one another) and developmental attributes (e.g., how the MTS membership changes over time) (Zaccaro et al., 2012). The developmental attributes reflect the fluidity of MTS, illustrating that the composition of, and relationships between, component teams will change and evolve over time in response to changing demands (Zaccaro et al., 2012). The additional structural considerations faced by MTS suggest that team processes may manifest differently to teams operating alone, as MTS teams must be able to manage inter- and intra- team processes in the midst of structural changes to the system over time (Zaccaro et al., 2012). Two processes where the effect of time may be most obvious are communication (a reciprocal process of sending and receiving information) and coordination (the enactment of behavioural and cognitive mechanisms that enable teams to synchronise their efforts to achieve goal related actions) (De Dreu & Weingart, 2003, LePine, Piccolo, Jackson, , Mathieu, & Saul, 2008).

Communication

Clear and open communication is essential to enable MTS to function effectively (Keyton, Ford & Smith, 2012). Establishing effective communication channels in MTS can be challenging due to a lack of familiarity amongst team members, an absence of common organisational language and the fact that teams may be geographically dispersed (Fodor & Flestea, 2016; Mishra, Allen & Pearman, 2011; Waugh, 2004). These challenges can result in a lack of clarity about what information needs sharing with who and when. This can reduce a MTS's ability to achieve low levels of *information opacity* (Luciano et al., 2018). According to Luciano et al., low levels of information opacity, as marked by the absence and ambiguity of information across component teams, enables MTS to adopt a common language, develop a shared understanding of the situation and establish effective communication channels which ultimately lead to better performance. This is supported by research linking effective communication with greater shared awareness (Salas, Prince, Baker & Shresetha, 1995), collaboration to achieve higher-level superordinate goals (Davison, Hollenbeck, Barnes, Sleesman, & Ilgen, 2012; Keyton et al., 2012) and meshing of skill sets to coordinate more effectively (Firth, Hollenbeck, Miles, Ilgen & Barnes, 2015). Conversely, when information opacity is high, activities of teams within the MTS remain relatively unknown to each other, which can lead to confusion and a lack of shared awareness.

Achieving low levels of information opacity depends on the challenges presented to the MTS at any one time. Emergencies are unpredictable and so certain teams (focal teams) may be more central than others to the achievement of MTS goals at any given point in time (Davison et al., 2012; Fodor & Flestea, 2016). This is referred to as the *fluidity of the structural configuration* of the MTS and it can lead to changes in how team members interact with one another over time (Luciano et al., 2018).Whilst this fluidity is necessary to allow focal teams to focus on salient tasks, it may lead to high levels of information opacity as team members fail to relay information to the remaining component teams. The response to the 2017 Manchester Arena terrorist incident illustrates this process. The demands placed on Police commanders to deal with the immediate aftermath of the explosion directly contributed to a lack of communication with inter-agency colleagues. This absence of clear communication across agencies led to a two-hour delay in the deployment of the Fire Service to the incident ground meaning their skills and expertise could be not be utilised until later on in the response (Kerslake, 2018).

One way to gain insight into the level of information opacity in a MTS is social network analysis. Social networks show the pattern of interactions between individuals and teams as measured through communication (Knoke & Yang, 2008). Of particular interest is the degree to which MTS are centralised; essentially the extent to which interactions within the network are concentrated within a small number of component teams (Wasserman & Fraust, 1994). Studies of MTS have identified beneficial effects of de-centralised structures, which facilitate low levels of information opacity by enabling information to be shared easily across the network, promoting shared awareness and improved performance (Brown & Miller, 2000; Lanaj, Hollenbeck, Ilgen, Barnes & Harmon, 2013; Schraagen & Van de Ven, 2011).

There is an absence of empirical research exploring how network structures change over time and under what constraints de-centralised structures can be achieved (Waring et al., 2019). Centralised networks with a single focal team may be unavoidable in the early phases of emergency response (e.g., when the incident is ongoing) to allow decisions to be taken quickly without the need for agreement across multiple agencies. When the initial urgency of the incident subsides, and the MTS shifts into the recovery phase (e.g., restoring business as usual) a less centralised system may be more easily achieved as a greater number of agencies work together to ensure communities and infrastructure are restored. In this Study we explore how communication networks, and thus information opacity, differ in a MTS at different points during a simulated incident response: incident ongoing, 48 hours after the incident (response phases) and three weeks after (recovery phase) and explore how this relates to changes in contextual demands.

RQ1: How will MTS communication networks change over time as the incident progresses from the response phase of the emergency (during an ongoing incident and soon after) to the recovery phase (returning to normality post-incident)?

Coordination

A second MTS processes that may be affected by time is coordination. For MTS, coordination is a complex process that encapsulates a multitude of behaviours that enable component teams to achieve both inter- and intra- team goals (Mathieu et al., 2018; Shuffler et al., 2015). Whilst essential for effective performance, coordination can be especially difficult in MTS when compared to conventional teams (Larsen, Nystad & Taylor, 2014). In conventional teams, coordination is underpinned by the development of transactive memory systems and team mental models: the knowledge possessed by different team members of who in the team knows what, which develops over time as teams become more familiar with one another (Austin, 2003; Smith-Jentsch, Kraiger, Cannon-Bowers, & Salas, 2009). For MTS, the development of transactive memory systems is made difficult due to the often-short periods that they work together. This is shown in studies that have attributed the link between a lack of role understanding and poor coordination in multi-agency teams to a lack of familiarity across agencies (Salmon et al., 2011; Waring et al., 2019)

According to Mathieu et al. (2001), MTS operate within a goal hierarchy, comprising proximal (intra-agency) and distal (inter-agency) goals. Conflicting goals across component teams can disrupt coordination and reduce performance (DeChurch & Marks, 2006; Luciano et al., 2018). This disruption is referred to as *work process dissonance* and reflects the extent to which component teams are focused on incongruent and independent tasks (Lucianio et al., 2018). The hierarchy of goals changes over time dependent on contextual demands, creating additional challenges as teams continually reappraise how actions are directed towards changing goals and priorities (DeShon Kozlowski, Schmidt, Milner & Wiechmann, 2004; Lucianio et al., 2018). The challenge for MTS is how to coordinate their efforts to successfully meet these changing priorities. For example, in the response phase, different agencies may be focused primarily on different goals: Ambulance and Fire commanders will be focused on extracting and treating casualties, whereas Police commanders will be focused on mitigating further threats and recovering evidence from the scene. However, in the weeks following the emergency, teams may be unified in their focus of restoring infrastructure and ensuring that business as usual can resume in the surrounding area ("recovery phase"). There is limited research examining how coordinating behaviours change over time (Davison et al., 2012). This means we have a limited understanding of how teams maintain effective coordination whilst goal hierarchies shift dependent on contextual demands (Luciano et al., 2018).

The MTS literature is relatively under-developed in its understanding of what constitutes coordination (Wijnmaalen, Voordijk & Rietjens, 2018). Some studies have approached coordination as a single measured behaviour (*see* Firth, Hollenbeck, Miles, Ilgen & Barnes, 2015). For example, Firth et al. measure MTS coordination as synchronised target selection by multiple component teams during a simulated task. However, this may fail to capture the complexities of coordination and how it changes over time. Rather than viewing coordination as a single behaviour, it may be better understood as a series of behaviours that can serve to promote or disrupt system and component team performance dependent on contextual demands (Rico et al., 2018). For example, sharing resources to support a focal agency might be indicative of effective coordination in the early phase of an emergency. However, as the incident progresses, joint decision-making across agencies may be needed to coordinate multi-agency action effectively. One way to understand what coordinating behaviours look like in MTS, and how they change over time, is to examine the detail of team interactions. Qualitative analysis of team interactions generates a deeper, fine-grained understanding of MTS behaviours than can be gained from quantitative approaches (Shuffler et al., 2015). Accordingly, the second aim of this Study was to identify verbal indicators of coordinating behaviours in the MTS response to an emergency, and, through qualitative analysis, explore how these behaviours change over time.

> *RQ2: What are the verbal indicators of coordinating behaviours in the MTS response to an emergency? How do these behaviours change as the incident progresses from the response phase of the emergency (during emergency) to the recovery phase (returning to normality post-incident)?*

Research Context

The research questions will be explored in the context of multi-agency emergency response in the U.K. In the U.K, an emergency is defined as any event that threatens the welfare, environment or security of the U.K, including acts of terrorism or war (Civil Contingencies Act, 2004). A 'major incident' is a specific type of emergency requiring the response of multiple agencies due to its multi-faceted nature (e.g., major road traffic accident, terrorist incident) (Cabinet Office, 2012). This response follows an Integrated Emergency Management (IEM) framework, where coordination is required at three levels: operational (bronze), tactical (silver) and strategic (gold) (Salmon et al., 2011). Operational commanders operate at the lowest level of command, responding on the ground to the immediate threat of the incident. At the tactical level, commanders from key agencies combine (e.g., Police, Fire, Ambulance) to ensure actions taken at the operational level are coordinated, cohesive and integrated. Tactical commanders prioritise resources and identify risks to notify strategic commanders. The Strategic Coordinating Group (SCG) take overall responsibility for the incident, gathering information from several sources to initiate a clear strategy that considers the possible longer-term and wider-reaching impacts of the emergency (McMaster & Baber, 2012).

This Study focuses on collecting data at the strategic level. SCGs are attended by the largest number of agencies, meaning they are suited to understanding larger MTS in which communication and coordination is expected to be difficult (Lanaj et al., 2013). Major incident response comprises two phases: response and recovery (Cabinet Office, 2012; Civil Contingencies Act, 2004). Response is defined as the decisions and actions required to deal with the immediate effects of an emergency, usually lasting no longer than a few days (e.g., identifying a reception centre for those caught up in the emergency). Recovery may last months or years, with decisions aimed at re-building and supporting communities following an emergency (e.g., restoring business as usual). The features that define each phase and the urgency by which decisions need to be made, are therefore different. Accordingly, we would expect to see differences in how teams communicate and coordinate dependent on the phase of the response.

Method

Participants

Participants attended a one-day table-top simulation event on the response to a terrorist incident as a training exercise. Participants were 30 strategic decision-makers and support staff (e.g., Loggists) from emergency response authorities, who, in the event of a real incident, would be likely to respond. Experience ranged from four months – 38 years (M = 16 years). To preserve anonymity, and in agreement with training coordinators from the Emergency Services, no other demographic data was collected. Three SCGs took place which were designed to emulate different time points following a terrorist incident: SCG1 (incident ongoing), SCG2 (48 hours after), and

SCG3 (3 weeks after). SCG1 and SCG2 were classed as response phases and SCG3 was classed as a recovery phase. The majority of participants took part in all three SCGs (n = 23). Variation in participation of the remaining members was reflective of how agencies would be represented in the event of a genuine incident: (i) one military representative only took part in SCG1; (ii) one Police representative only took part in SCG2 and SCG3; (ii) an additional member of Coroner's office took part in SCG2; (iii) no members of the Coroner's office took part in SCG3 and (iv) an additional representative of local business owners took part in SCG3. The number of team members present from each agency in each SCG is shown in Table 1. Ethical approval for this Study was obtained from the Faculty of Science and Technology Research Ethics Committee (FST16088), Lancaster University.

Table 1. Agencies represented at each time point

Agencies	SCG1	SCG 2	SCG 3
Police	5	5	4
Ambulance and NHS	6	6	6
Fire and Rescue	1	1	1
Military	2	1	1
Coroner's Office	1	2	0
Government	2	2	2
Local Council	8	8	8
Local Business Owner	2	2	3
Red Cross	1	1	1
Total	**28**	**28**	**26**

Procedure and Exercise Scenario

Before the start of the exercise, participants were informed that psychologists were present at the training exercise and that research on teamwork during major incident response was taking place[1]. Participants were asked if they were willing to be audio-recorded during the simulation and were invited to ask any questions. All participants agreed and once informed consent was obtained, the exercise commenced. The exercise was based on the response to an armed terrorist attack in a busy shopping centre on a Saturday evening in the run up to Christmas. Data were collected from three SCGs, each dealing with a different time point in the incident: (i) incident ongoing; (ii) 48 hours after and (iii) three weeks after. Accordingly, the tasks completed during each SCG varied dependent on changing demands (*see* Appendix 1 for a detailed overview of tasks). For example, during SCG1 participants were focused on responding to the immediate threat of the incident and identifying victims whereas during SCG3 the focus turned to business continuity and establishing a long-term ministerial recovery group. Collecting data at multiple time points was important to illustrate the transition from response phases of the incident (SCG1, SCG2) to recovery phase (SCG3), and to capture the dynamic nature of MTS responding in critical, uncertain environments (Aiken & Hanges, 2012). For each SCG, participants were seated around a large table and received information about the incident in written and audio form. Each SCG was tested in sequence with a short refreshment break in between. SCG1 lasted 81 minutes, SCG2 lasted 49 minutes, and SCG3 lasted 65 minutes.

The exercise was run as an immersive table-top simulation, developed by the training team of the Local Resilience Forum. The exercise was designed by training coordinators to ensure that the response required collaboration across agencies, testing the principles of the Joint Decision Model (the standardised protocol for multi-agency response to emergencies in the U.K; JESIP, 2013). Immersive simulations allow

---

[1] Access to the training exercise was granted due to pre-existing relationships with the exercise organisers. In exchange for access to the data, the team prepared a debrief and short summary report evaluating the response for practitioners.

researchers to collect rich data with expert practitioners, re-creating the challenges of the organisational context in a safe and ethically appropriate context (Alison et al., 2013). Simulation studies are vital to advance understanding of teamwork during major incident response (*see* Alison et al., 2015b; Wilkinson, Cohen-Hatton & Honey, 2019), especially in instances where direct observation and traditional laboratory experiments are not possible due to the security sensitive nature of the research (e.g., the response to a terrorist incident).

Data Collection and Analysis

Audio recording. Five Dictaphones were placed on the table in the middle of the testing room to audio-record the communications between team members during each SCG. Recordings were later transcribed verbatim.

Social network analyses. Communication data for each SCG were analysed separately using social network analyses (SNA). SNA are useful for providing a visual representation of the social dynamics of a team, analysing the prevalence and frequency of interactions between team members and across component teams within a MTS (Driskell & Mullen, 2005; Fodor & Flestea, 2016). In common with prior research, we used degree centrality (centralisation) to explore information opacity in MTS (Fodor & Flestea, 2016). Centralisation was calculated using Freeman's (1978) formula and reflects the extent to which interactions are concentrated within a small number of component teams (Wasserman & Fraust, 1994). A completely centralised network would have a value of 1 and resemble a star shape as all communications pass through one centralised node; whereas a decentralised network would have a value of 0.

As the networks were directional (i.e., they showed the direction of interactions between team members), we also computed in-degree and out-degree network centrality. High out-degree centrality suggests a small number of agencies are sending out information to other nodes; whereas high in-degree centrality suggests most communications in the network are directed towards a small number of nodes (Scott, 2000). We also calculated network density to show the proportion of network ties against the total number of possible ties (Wasserman & Fraust, 1994). Finally, to show how individual contributions changed across the SCGs, we calculated itemised indegree and outdegree centrality.

A communication matrix for each SCG was created from the transcribed data. The matrices were loaded into R and centralisation, in and out-degree centrality and density were calculated using R package "sna" (Butts, 2008). Gephi was used to create a visualisation of the communication networks between team members for each SCG (*see* Figures 1-3). Each node in the network represents a team member, and the size of the node equates to its degree centrality within the network. The arrows indicate the direction of communication between the speaker and the receiver and the thickness of the line indicates the frequency of communication between them. Agencies within the MTS are differentiated by the colours of the nodes.

Thematic analyses. A thematic qualitative analysis was conducted on the transcribed data to identify verbal indicators of coordinating behaviours (Braun & Clarke, 2006; Nowell, Norris, White & Moules, 2017). Using communication to code for coordination is justified as explicit coordination cannot occur without the communication of component teams as it reflects the verbal interactions required to align the goals and priorities of different component teams (DeChurch & Marks, 2006; Espinosa, Lerch & Kraut, 2004). Coordinating behaviours were regarded as behaviours that enabled team members to synchronise their efforts to achieve goal related outcomes (De Dreu & Weingart, 2003; Marks, Mathieu & Zaccaro, 2001). Analyses therefore included identifying and coding behaviours that enabled team members to synchronise their efforts to achieve goal related outcomes (positive behaviours) and those that disrupted their efforts (negative behaviours). In taking this approach, we identified behaviours relevant to task completion (a process measure), rather than taking an objective measure of team performance (an outcome measure). Team performance measures can be difficult to devise or interpret in complex, extreme environments in which decisions often have no right or wrong answer (Alison et al., 2015a; Mathieu Maynard, Rapp, & Gilson, 2008).

To code the data, we followed Braun and Clarke's (2006) six-phased method, combing inductive and deductive analysis (Fereday & Muir-Cochrane, 2006). Following familiarisation with the data, we drew on prior literature and theory to guide the analysis and develop the initial codes for the codebook (deductive). For instance, *joint decision-making* and *sharing task related information* were included in the initial codebook. Both are core principles of coordinated working in the emergency services and are identified in prior research as important for MTS coordination (JESIP, 2013; Davison et al., 2012; Wijnmaalen et al., 2018). We then searched for additional codes that emerged from the data (inductive). Higher order themes were derived from the initial second order codes within the coding dictionary (Appendix 2). The themes were reviewed and defined in detail and the transcripts were coded for a second time to ensure intra-rater reliability (k =.74) A second coder was then trained to apply the coding dictionary to identify higher order themes in each of the SCG transcripts (20%). An acceptable level of inter-rater reliability was obtained (k = .68) (Everitt, 1996).

A frequency count of the behaviours was calculated to compare coordinating behaviours across different time points. To account for the different lengths of each SCG phase, proportional counts of behaviours were calculated as a percentage of the overall number of behaviours in each SCG, and these values were used in the main analyses.

Results

Communication networks of the MTS across time

The results of the SNA showed that the total number of interactions (weighted edge) between team members reduced from SCG1 to SCG2 to SCG3, however, the number of unique interactions (edges) remained similar across SCGs (Table 2). The results of the itemised in-degree and out-degree scores in Table 3 suggests this reflects a greater number of in-going and out-going communications by the chair (in this instance the Police) in SCG1, rather than an increase in communication across component teams. This finding is consistent with the scores of network density which remained consistently low throughout.

Table 2. SNA statistics for each SCG

SNA statistics	SCG 1(*n*=28)	SCG 2(*n*=29)	SCG 3(*n*=26)
Total number of weighted edges (overall number of interactions)	544	419	327
Total number of edges (number of unique interactions)	58	60	57
Density	.077	.073	.087
Centralisation	.86	.77	.76
Degree centrality (in)	.73	.57	.55
Degree centrality (out)	.92	.92	.91

Centralisation was used to explore how information opacity changed during the response (Table 2). The results showed that SCG1 was highly centralised and reliant upon a single team member (i.e., the chair), whereas SCG2 and SCG3 were less centralised. This suggests lower levels of information opacity in the latter SCGs as other agencies became more involved in discussions. Whilst out-degree centrality remained high throughout indicating a small number of agencies were responsible for providing information to other team members, in-degree centrality decreased. This decrease is evident of a greater number of team members being consulted over time and is consistent with the centrality data that the Chair became less focal for team discussions over time. It is also consistent with the changes in the itemised in degree and out degree scores (Table 3) which show contributions from a greater number of team members over time. Findings are visualised in the network plots, as can be seen in Figures 1-3.

Table 3. Itemised indegree and outdegree scores for each participant in each SCG

	SCG1		SCG2		SCG3	
	In	**Out**	**In**	**Out**	**In**	**Out**
Police Gold	22	27	18	28	4	4
Government Liaison	1	1	1	1	1	1
Fire and Rescue Gold	1	1	1	1	1	1
Red Cross	2	2	3	3	1	1
Special Branch Police	1	1	1	1	-	-
Council D Gold	1	1	1	1	1	1
Council C Gold	1	1	1	1	1	1
Council B Gold	1	1	1	1	4	4
Council A Gold	2	3	3	2	16	25
Military Liaison	1	1	-	-	-	-
Regional Military	1	1	3	2	1	1
Shopping centre Operations	2	1	1	0	1	0
Shopping centre Director	1	1	2	3	5	5
NHS Advisor	1	1	1	0	1	0
Ambulance Gold	2	1	1	1	1	1
NHS Gold	4	4	3	3	3	2
Coroner 1	1	2	2	2	-	-
Senior Investigating Officer	3	2	3	3	2	1
Tactical Firearms Commander	1	1	2	3	1	1
Media for Police	1	1	2	2	3	3
Support for NHS	1	0	1	0	1	0
Support for Ambulance	1	1	1	0	1	0
Loggist for Ambulance	1	0	1	0	1	0
Support for Council C	1	0	1	0	1	0
Support for Council B	1	0	1	0	1	0
Support for Council A	1	2	1	0	1	0
Support for Council D	1	0	1	0	1	0
COBR Call	1	0	1	1	1	1
British Land	-	-	1	1	2	4
Coroner 2	-	-	1	1	-	-

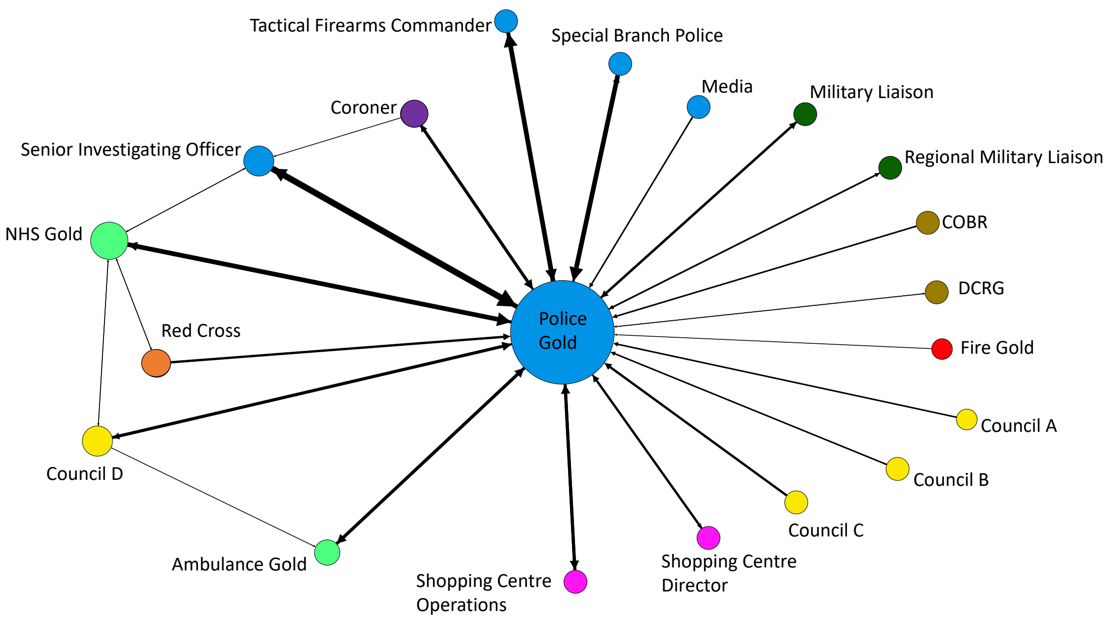


Figure 1. Network plot of SCG1[2]

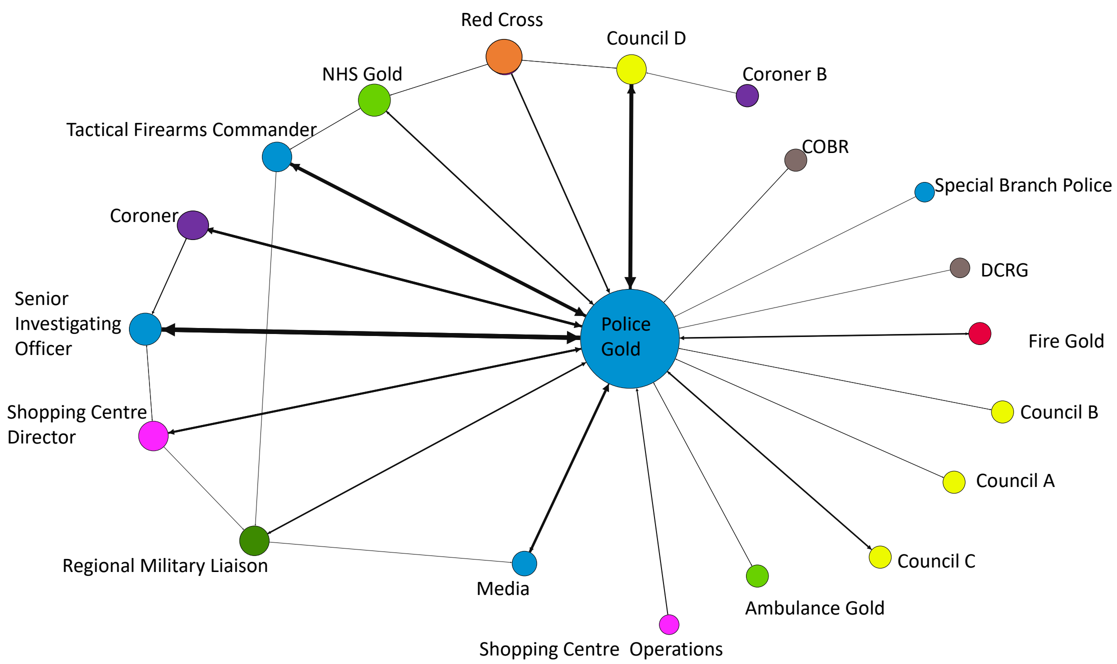


Figure 2. Network plot of SCG2

[2] Please note the size of each node equates to its degree centrality within the network and the colour of the node distinguishes between different agencies within the MTS. The arrows indicate the direction of communication between the speaker and the receiver and the thickness of the line indicates the frequency of communication between them.

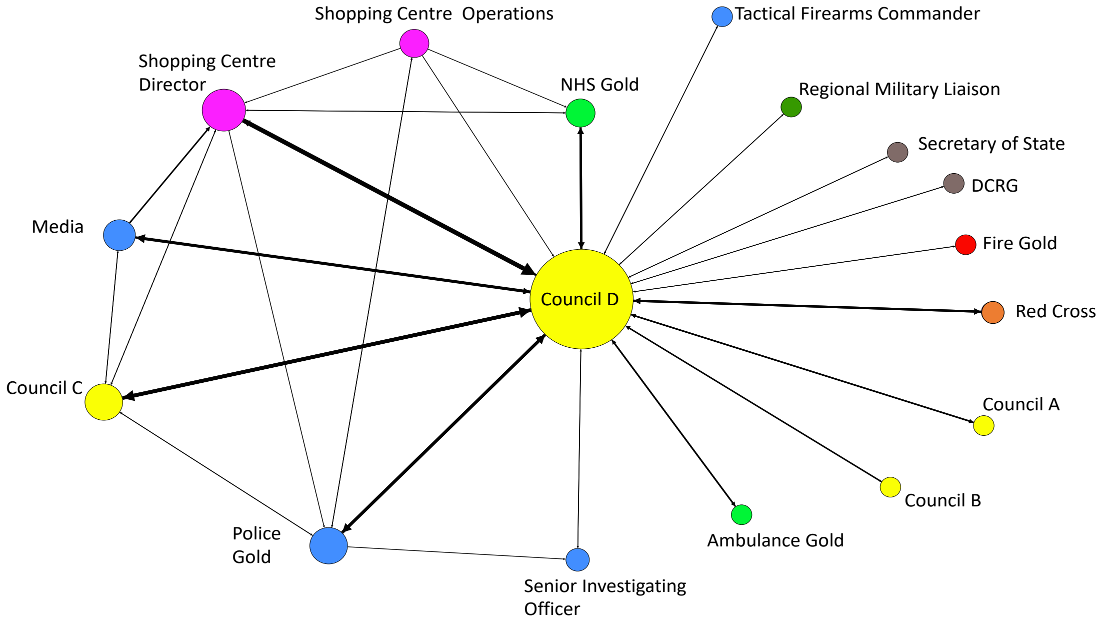


Figure 3. Network plot of SCG3

Coordinating behaviours in MTS interactions across time

The second research question sought to: (i) identify the coordinating behaviours in MTS; and (ii) show the frequency of these behaviours at different time points (*see* Appendix 2). A thematic analysis of communications during the SCGs identified three positive and three negative team behaviours (*see* Table 4). Positive behaviours were: *joint decision-making*, *sharing resources*, and *sharing task related information*. Negative behaviours were: *role uncertainty; decision uncertainty and conflicting priorities.*

Table 4. Proportion of positive and negative coordinating behaviours across each SCG

Behaviour	Description	SCG 1	SCG 2	SCG 3
Positive		**28 (75.6%)**	**23(84.2%)**	**18 (94.7%)**
Joint decision-making	Instances in which the team actively worked together to implement a decision	8 (21.6%)	11 (40.3%)	8 (42.1%)
Sharing resources	Instance in which team members offered resources to assist other agencies (e.g., staff, roadblocks).	6 (16.2%)	3 (11.0%)	2 (10.5%)
Sharing task related information	Instance in which team members actively sought to share agency specific information and improve shared situational awareness.	14 (37.8%)	9 (32.9%)	8 (42.1%)
Negative		**9(24.3%)**	**4 (14.8%)**	**1 (5.3%)**
Role uncertainties	Confusion about one's own role and the role of others.	0 (0.00%)	1 (3.7%)	1 (5.3%)
Decision uncertainties	Instances in which misinformation led to confusion and discussions were delayed due to indecision.	6 (16.2%)	1 (3.7%)	0(0.00%)
Conflicting priorities	Instance in which team members attempted to re-orient the conversation towards intra-agency priorities.	3 (8.1%)	2 (7.4%)	0(0.00%)

Joint decision-making. *Joint decision-making* was coded as any instance in which the MTS actively worked together to implement a decision. An example of this occurred in SCG1 when team members were discussing how to quickly move casualties from the incident site to hospitals within the immediate vicinity of the shopping centre. NHS director referred directly to Police Gold to ensure that hospitals were identified that were secure and easily accessible *"we will need [to] disperse casualties outside of "Location A" so we need [to] ask for support to secure the hospitals in "Location B" and "Location C" in order to be able to take casualties there so we would ask for the support of Police forces in those areas" (SCG1, NHS Gold).* In addition, when the team were making a decision regarding their media strategy in SCG1, the chair encouraged team members to act in a coordinated manner: *"...bring back to this group any intended* *messages that are going out so that we don't have messages sent out in isolation that aren't coordinated"(SCG1, Police Chair).*

Sharing resources. *Sharing resources* was coded as any instance in which agencies offered resources to assist other agencies within the MTS. An example of this was evidenced during SCG1, when the Government Liaison Officer offered support to Police in the early stages on the incident in contacting ministers and liaising with Military officials: *"In terms of government support, I'm happy to facilitate anything in terms of speeding up the military assistance and contacting ministers*" *(SCG1, Government Liaison Officer*). A further example was evident during SCG2 when a discussion took place involving Council A about how to direct individuals wanting to make charitable donations. The Red Cross offered to help coordinate a central fund for the victims of the incident in conjunction with local councils: *"yes that's very much something we will be able to offer support on" (SCG2, Red Cross).*

Sharing task related information. Instances in which team members actively sought to share agency specific information and improve shared situational awareness were coded as *sharing task related information*. In each SCG, the chair encouraged team members to share information about their current situation: *"...an update on the current situation and any briefing from any individual agencies on where we are" (SCG3, Council D Chair).* In addition, the chair sometimes requested an explanation of agency specific acronyms, for example the terminology used to describe the severity of casualty cases: *"if a health colleague could just explain the P2, P3 and P1, just so we are all aware of the terminology please" (SCG1, Police Chair).*

Role uncertainties. *Role uncertainties* were coded if there was confusion about one's own role and/or the role of others. This can be exemplified in SCG2 and SCG3 in a dispute between NHS representatives and the Red Cross over who takes primacy in providing psychological support to those involved in the incident. During SCG3, the NHS are required to state their authority in a scenario in which the Red Cross imply setting a strategy of support: *"just to be clear on that... we will draw on colleagues from local authorities, but in terms of returning to normal, that responsibility sits within the NHS" (SG3, NHS Gold).*

Decision uncertainties. Instances when decision-making lacked clarity and discussions were delayed due to indecision were coded as *decision uncertainties*. An important example of this was during SCG1 when there was a lack of decisive action about what to do with the 30,000 individuals who had self-evacuated the site of the incident (a large indoor shopping centre). The issue was raised on multiple occasions by the Head of Operations of the shopping centre, without an adequate response: *"There is a high demand for information, that we are not really able to respond to at this moment in time" (SCG1, Head of Operations)*. Another instance of decision uncertainty occurred during SCG2, when discussing an appropriate place to host a memorial site. Local council members identified a site where the memorial should be established, however this information was ignored by others, leading to a lengthy debate about other possible sites: *"I'm just conscious that we need to deal with this issue of a memorial, team. So it's… [location]. That's the scene of the memorial?" (SCG2, Police Chair).*

Conflicting priorities. When team members attempted to re-orient the conversation towards intra-agency priorities this was coded as *conflicting priorities*. During SCG1 there were several moments when discussions began to turn to intra-team tactical matters. The chair was then forced to re-direct conversations to the strategic issues at hand: *"Folks, what I don't intend to do is to get tactically distracted as the* *updates come in. We have a tactical commander at this core who is taking care of that" (SCG1, Police Chair).*

As well as identifying coordinating behaviours, we wanted to explore how the frequency of these behaviours changed over the course of the response. Figure 4 shows the proportion of positive behaviours increased across the three SCGs, from 75.6% to 94.7%; whilst the proportion of negative behaviours decreased from 24.3% to 5.3%. A chi-square test of independence indicated the differences in behaviours were significant, χ^2 (2, n = 299) = 14.40, p = .001. Positive behaviours consistently outweighed negative behaviours across the three SCGs, with the prevalence of *sharing task related information* and *sharing resources* remaining relatively consistent throughout, and the prevalence of *joint decision-making* increasing from 22% in SCG1 to 42% in SCG3. The decrease in overall negative behaviours across SCGs was due to reductions in *decision uncertainties* (from 16.2 % in SCG1 to 0.0 % in SCG3) and *conflicting priorities* (from 8.1 % in SCG1 to 0.0% in SCG3). In contrast there was a slight increase in *role uncertainty*, from 0.0% in SCG1 to 5.3% in SCG3.

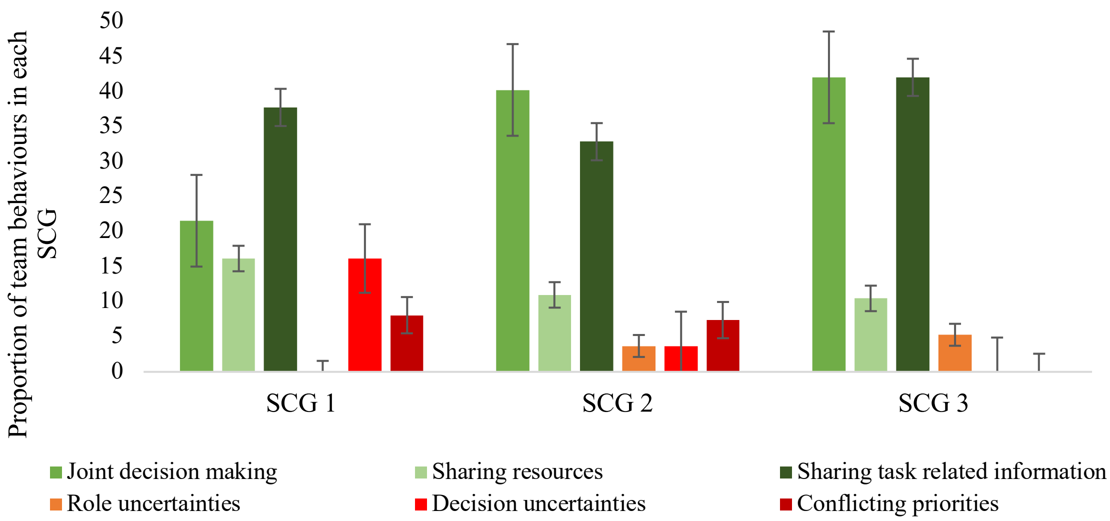


Figure 4. Changes in positive and negative team behaviours in each SCG

Discussion

This Study explored changes in communication and coordination in a MTS responding to an emergency. The first aim was to explore how the network of communications changed across each phase of the response. It was found that whilst SCG1 was highly centralised, with communications dominated by the Police, SCGs2 and 3 were more decentralised, with a greater number of agencies involved in discussions. This is interesting as SCGs1 and 2 were classed as 'response' phases by the authorities (during an ongoing incident and soon after) and thus were expected to be more similar when compared to SCG3, classed as 'recovery' (returning to normality post-incident). Second, we sought to uncover a more detailed understanding of what coordination looks like in MTS and how these coordinating behaviours changed across time. Three positive and three negative behaviours were identified. It was found that positive coordinating behaviours increased over the three timepoints of the incident and negative coordinating behaviours decreased. Taken together, these findings suggest that the frequency of positive coordinating behaviours increased as the incident progressed, and that this coincided with the MTS network becoming less centralised.

Findings from this Study demonstrated that, despite communication networks remaining relatively centralised throughout, centralisation decreased as time from the incident increased. Our results demonstrated that the involvement of additional agencies increased in SCGs 2 and 3 in comparison to SCG1, and this was evidenced by a decrease in network centrality and in-degree centrality. This suggests that information was shared from, and relevant to, a greater number of agencies in the latter phases of the response, resulting in lower levels of information opacity. The change in communication networks over time calls attention to the importance of time when examining communication network structures in MTS.

The phase of the response likely played a role in this finding. During the initial response to the incident (SCG1) agencies were dealing with a live firearms incident. Procedural guidelines dictate that when firearms are present, the over-arching responsibility and direction of the “response phase” rests with the Police (Cabinet Office, 2012). Findings in the network structure for SCG1 suggest that this occurred in the current Study. This finding provides empirical support for the *fluidity of the structural configuration* of MTS, in which the relative importance of a component team (in our example the Police) may be more central to the achievement of system level goals at a given point in time, dependent on task demands (Davison et al., 2012; Luciano et al., 2018). As suggested in the introduction, this could lead to high levels of *information opacity,* as the dominant team focuses on the task at hand, prioritising this over maintaining clear communication channels with the remaining component teams. Our results support this claim, as evidenced by missed information from inter-agency partners in SCG1 delaying decision making on a key issue: the evacuation of members of the public from the Shopping Centre. The Chair (Police Gold commander) was focused on delivering the overall strategy of the response and failed to attend to important information provided by inter-agency partners.

Once the immediate threat of the live incident had subsided (SCG2 and SCG3), the networks became less centralised. Interestingly, whilst SCG2 was still classified as the “response phase”, its communication network was structurally similar to SCG3 (“recovery phase”) than it was to SCG1. This suggests contextual demands present in the environment (i.e., void of a live ongoing incident) drove changes in the network structures, as opposed to teams following procedural guidelines. Similar findings were noted by Choi and Brower (2006), who found perceived network structures during emergencies were different to those prescribed in emergency management plans.

Together with Choi and Brower, our results suggest the network structures outlined in emergency guidelines do not reflect how real-world network structures change over time. Policy makers should consider reframing procedural guidelines to reflect the behaviours of responders during crises which would better prepare teams for emergency incidents.

The second aim of this Study was to identify positive and negative coordinating behaviours and how these change over time. Generating a deeper understanding of *how* coordinating behaviours occur in MTS is important for developing insights into how to improve MTS effectiveness (Mathieu et al., 2018). Existing work focuses on single measured behaviours (*see* Firth et al., 2015), which may be limited in representing the complexities of coordination in MTS. In this Study, we focused on verbal indicators of coordination, and as such, we captured instances of explicit coordination, as opposed to implicit coordination (e.g., team mental models) (Rico et al., 2018). Positive team behaviours were coded as *joint decision-making, sharing resources, and sharing task related information*. The results show that positive coordinating behaviours increased over time but were also consistently high across the three SCGs in comparison to negative coordinating behaviours.

During emergencies, *joint decision-making* (JDM) is important to enable teams to act in a timely manner (Chen et al., 2008), whilst ensuring that decisions do not compromise the policies or procedures of other component teams. It is possible that JDM was lower earlier in the response due to the differential “power distribution” of component teams caused by the *fluidity of the structural configuration* of the MTS (Luciano et al., 2018; Zaccaro et al., 2012). As mentioned earlier, it is likely that as SCG1 involved a live firearms incident, JDM was lower because of the higher power distribution residing with the police and thus leaving less room for inter-agency decision-making. When there is reduced JDM and a high level of centralisation, there is a risk of key agencies becoming overloaded. Another risk is that the over-involvement of a single agency may lead to a loss of expertise within the MTS, because other, less central, agencies participate less in decision-making. One of the main benefits to MTS is the diversity of skills from different agencies and this benefit should not be over-looked despite the urgency of an unfolding incident (*see* Kerslake, 2018). This echoes the findings in the Pollock report on interoperability in U.K emergency management (2017) which highlighted how, despite efforts to increase joint decision-making across agencies, many key decisions were still being taken by a single organisation. Our finding that increases in *joint decision-making* coincided with less centralised networks suggests implementing de-centralised communications networks may increase collaborative decision-making across agencies.

The second positive behaviour identified was *sharing resources*. During emergencies sharing resources is important, as component teams may find themselves overloaded and requiring assistance of other agencies (Power, 2018). Slightly fewer instances of *sharing resources* occurred in SCG2 and SCG3, though still higher than the negative behaviours. This may have occurred as offering support and *sharing resources* was more important in the early phase of the response as the team put systems and protocols in place. In the latter phases of the response, many of these systems and protocols were in place and so communications were less resource focused. The final positive behaviour is *sharing task related information* (STRI). STRI is vital in emergency response as it allows component teams to reach a collective understanding of the situation, enabling them to coordinate effectively (Salmon et al., 2011; Van der Haar, Segers, Jehn & Van den Bossche, 2015). While STRI remained high throughout each SCG, the results of the SNA show that in the latter stages, the information shared came from and was discussed by a larger number of agencies.

We identified three negative behaviours relating to coordination: *role uncertainties, decision uncertainties, and conflicting priorities*. *Decision uncertainties* were coded as a delay in decision-making due to missed information or indecision. These were more prevalent during the response phase, possibly due to a higher level of uncertainty and ambiguity in the task. Notably, in SCG1, the MTS continually failed to decide what to do with the 30,000 individuals who self-evacuated from the shopping centre. Responders raised the issue numerous times but failed to implement a decision. The prevalence of *decision uncertainties* when deciding how to evacuate the general public, supports the notion that when task uncertainty is high in MTS, inter-team processing is challenged (Luciano et al., 2018). Empirical evidence suggests uncertainties are alleviated if communications are increased across agencies (Phelps, 2010). Our finding that *decision uncertainties* became less prominent as networks became less centralised supports this. When networks are less centralised, information can be transferred more easily across agencies (Wasserman & Faust, 1994), which may increase shared awareness and reduce uncertainty around the task.

Conflicting priorities reflected team members re-orientating discussions to suit their own intra-agency priorities. The higher prevalence of *conflicting priorities* during earlier SCGs reflects differences between the goals of component teams. When teams have conflicting priorities, this disrupts coordination and reduces performance (DeChurch & Marks, 2006). For example, during SCG2, the Police were trying to establish victim numbers with the Ambulance service, whilst council members were focused on re-building community cohesion. This provides empirical support for Luciano et al.'s (2018) meso-theory of MTS functioning, in which MTS can be *differentiated* by the level of work process dissonance in the teams (i.e., the extent to which task completion is independent and incongruent across teams). As indicated by increased instances of *conflicting priorities* and *decision uncertainties* earlier on in the incident, when work processes dissonance is high (i.e., team members focused on different tasks), this can impede team processes.

Whilst cumulatively negative behaviours decreased over time, this was not true for *role uncertainties*, which only occurred in the latter two SCGs, albeit in low frequency. Role uncertainty has previously been identified as a hindrance to coordination in MTS during major incident response, leading to poor understanding about opportunities for collaboration and unrealistic expectations between team members (Power & Alison, 2017a; Waring et al., 2019). The presence *of role uncertainties* in the latter phases of the response may be linked to the network becoming less centralised over time, with a greater number of agencies (e.g., the local council) involved in discussions, increasing the likelihood of team members' demonstrating a lack of understanding of one another's roles. Despite researchers theorising that additional component teams can create further challenges in MTS (Luciano et al., 2018), we did not find that the *role uncertainties* dominated discussions or reduced the ability of teams to continue making joint decisions (as shown by the higher frequency of JDM in the latter SCGs). Results therefore suggest that the involvement of additional component teams, whilst increasing role uncertainty, did not impede the ability of team members to work together.

Taken together findings are indicative that coordinating behaviours improved as the network of communications became less centralised and lower levels of information opacity were achieved. Primarily, this provides empirical support for the link between effective communication and coordination in MTS (Davison et al., 2012;

Firth et al., 2015). Whilst the direction and causality of this relationship cannot be determined here due to a small sample size (*N* = 3 simulations), it is in line with prior research that de-centralised networks benefit MTS during complex tasks (Schargen & Van de Ven, 2011; Brown & Miller, 2000; Lanaj et al., 2013). Findings extend prior research by demonstrating how communication and coordination change over time in the MTS response to an emergency. Showing how time-dependent factors (e.g., changing contextual demands) may have contributed to these changes, emphasises the importance of studying MTS across time.

Practical implications

Mapping communication networks over time illustrated the disconnect between different agencies during the response and demonstrated an over-reliance on the Chair to manage the flow of communications. This highlights the demand placed on the Chair: they are required to simultaneously manage communications across the MTS and take the lead on key decisions (Kerslake, 2018; Waring, Moran & Page, 2020). One way to lessen the load on the Chair and increase the connectedness across component teams might be to assign a "boundary spanner" in the response to crises. Boundary spanners are designated team members, tasked with ensuring that information is relayed, and actions are coordinated across different component teams (Carter, 2014; Chaffin et al., 2017). Accordingly, an effective boundary spanner can ensure low levels of information opacity are achieved throughout the response by increasing shared awareness across agencies. Further research is needed to explore if assigning this role to an individual would create more space for shared communication across the MTS and free up those in leadership to focus on decision-making and coordinated action.

The present Study involved an in-depth qualitative examination of the behaviours that constitute coordination in MTS. Findings can be used to inform the development of targeted training programmes to improve teamwork by specifically aiming to improve positive and reduce negative coordinating behaviours. For example, *role uncertainties* might be reduced by training component teams to clearly understand the responsibilities and skillsets of their inter-agency partners. Role clarification training has successfully improved performance in medical teams and future research might trial this intervention with multi-agency emergency response teams (Klein, DiazGranados, Salas Burke, Lyons & Goodwin, 2009; Salas, DiazGranados, Weaver, & King, 2008). Relatedly, if roles and responsibilities are clearly defined, this may better equip MTS to adopt de-centralised structures earlier on in the response as there will be a greater awareness of who to go to information (Luciano et al., 2018; Power & Alison, 2017b).

Exploring how coordinating behaviours changed across time also identified which behaviours were more likely to occur in different phases of the response. This could have important implications for where to focus efforts to improve coordination in MTS. For instance, *decision uncertainties* were markedly higher in the early phase of the response in which the incident was ongoing. As we found that *decision uncertainties* were in part attributed to missed information across agencies, it may be possible to reduce instances of decision uncertainties by implementing de-centralised structures in early phases of incident response. This would increase inter-agency communication and improve shared awareness.

Specific to response in the U.K, the findings of the network analysis suggest that the current response/recovery classification might not fit the nuanced context of an emergency. Unique to the present research was the inclusion of three different time points following the incident, two categorised as "response" (SCG1, SCG2) and one as "recovery" (SCG3). Both SCG1 and SCG2 were classed as 'response' and so we expected them to have similar network structures. However, we found that communication during SCG2 (response) was similar to SCG3 (recovery). It is possible that the 'response' phase should be limited to the immediate active phase of an emergency. In the current Study, and as per U.K Government guidelines, 'response' was still active 48 hours after the incident. Although there were still major challenges to manage, the active live risk was reduced as the terrorists had been killed. It might be possible to introduce a 'warm response' phase that acknowledges this reduction in immediate risk as time pressure has reduced from minutes to hours. This adjustment in how the SCG is framed could empower additional agencies to become more involved earlier on in the incident. This would lessen the load on other agencies such as the Police and increase opportunities for collaborative working across agencies.

Limitations and future directions

This research involved the qualitative and quantitative analysis of 195 minutes of rich data from 30 strategic responders. We acknowledge that the specificity of the context may limit generalisability and further research will be necessary to establish if findings apply to other MTS responding in similar high-stake settings. This Study does however represent one of few examples of empirical research that has sought to explore team processes in MTS and how they change over time, outside a laboratory setting. It is hoped this will encourage other researchers to begin marrying applied and theoretical approaches to studying MTS to move beyond theoretical frameworks and begin establishing how multi-agency teams perform *in situ.*

Despite noting a co-occurrence of less centralised communication structures and improved coordinating behaviours across time, our small sample size (*N* = 3 simulations) meant that we were unable to test the causal link between the two. It is therefore important to acknowledge confounding factors that may have been responsible for the changes. Firstly, each SCG had a different Chair. Whilst it is not uncommon to see different SCG Chairs in practice, it is possible that changes in communication and coordination observed in this Study were a result of changes in leadership style (*see* DeChurch & Marks, 2006). Future research should measure leadership behaviours in MTS to understand what relationship, if any, this has with how teams manage different phases of emergency response. Secondly, each simulated SCG took place on the same day and in quick succession, with a short refreshment break in between. Compressing time in simulation studies of emergency responders is important to increase immersion and ensure participants remain engaged with the simulated incident (Alison et al., 2013). Nonetheless, it is possible that the changes in behaviours we saw in the latter phases of the response were due to increased familiarity amongst team members improving team processes over the course of the day. However, any effect of familiarity on team behaviours is reminiscent of what we would expect to see in the response to genuine incidents as SCGs happen daily following major incidents. This effect of familiarity might be examined in future studies of MTS, to better understand how knowledge and experience of working with inter-agency partners impacts inter-team team processes.

Conclusion

This Study explored team processes in the MTS response to an emergency. We found that as time from the incident increased, the communication networks became less centralised, indicating that more agencies were involved in the decision-making. It was also found that as time from the incident increased, positive coordinating behaviours increased, whereas negative behaviours decreased. Findings demonstrated how team processes changed dependent on contextual demands and emphasise the relationship between communication and coordination in MTS. Implementing decentralised structures in the early phases of incident response may lessen the load on focal agencies (i.e., the Police), and increase opportunities to share information and coordinate decisions across inter-agency partners. Further research is needed to establish if the relationship between improved communication and coordination is causal and to identify if similar patterns of behaviour can be identified in MTS responding in analogous settings.

Chapter IV

Multi-Team Systems in Crises:

Does Familiarity Matter?

Chapter III showed the importance of time in understanding teamwork in a multi-team system (MTS) responding to an emergency incident. As time from the incident increased, communication networks became less centralised and this coincided with improved coordinating behaviours. Findings suggest in the early phases of response, commanders from key agencies (typically the Police) can become overloaded and a lack of involvement from other agencies can lead to failures in using a full range of expertise, resulting in less effective teamwork. This is in line with the findings of a recent Government report reviewing the response to the 2018 Manchester Arena terrorist attack (Kerslake, 2018).

The MTS for the Study in *Chapter III* was brought together quickly to deal with an unfolding emergency incident, meaning that responders from different agencies had little previous experience of working with one another. Evidence from the conventional team literature suggests that team tenure, or the level of familiarity amongst team members, can have a positive impact on team functioning (Espinosa, Slaughter, Kraut & Herbsleb, 2007; Huckman, Staats & Upton, 2009). Further, researchers have theorised that a key reason for continued challenges in MTS teamwork is a lack of familiarity amongst multi-agency colleagues (Fodor & Flestea, 2016; Waring, Alison, Shortland & Humann, 2019). This suggests had the MTS members in *Chapter III* been familiar with one another, we may have seen a higher frequency of positive behaviours/increased communications across inter-agency partners. Indeed, in discussion with commanders during the evaluation of the exercise, one participant reported that the response would have been improved with *"prior knowledge and familiarisation of working with the colleagues around the table."*

In light of this, *Chapter IV* presents one of the first studies to examine how familiarity impacts team behaviours in MTS. Given the context in which MTS operate, it may not be possible to ensure all team members are familiar with one another on an individual level. However, establishing if familiarity has a positive effect on behaviours in MTS may have implications for the importance of inter-agency training. Accordingly, the purpose of this Chapter was to test if familiarity was associated with improved team behaviours in the MTS response to a simulated emergency incident. In addition, we unpacked this relationship to test if the effect of familiarity on team behaviours occurred through shared goals and communication structures. Audio and video-recorded data were collected from 22 teams (n = 11 familiar, n = 11 unfamiliar) taking part in a simulated response to a terrorist incident. Each simulation was audio and video recorded to compare coordinating behaviours, shared goals and communication structures in familiar and unfamiliar teams.

The design of this study differs from the study in *Chapter III* in two important ways. First, we examined responses at one time only. We did this as the central focus was comparing familiar and unfamiliar teams, as opposed to exploring how behaviours changed over time. Second, we drew on a student population. We did this to ensure that familiarity could be manipulated across groups. As the simulations in *Chapter III* were part of a formal training exercise, our involvement was limited to observation and variables could not be manipulated for research purposes. Further, using a student sample ensured there was an adequate number of teams to compare the effects of familiarity and to begin developing a more complex understanding of the causal relationships between team processes. Although the co-occurrence of communication network changes and coordinating behaviours were noted in *Chapter III,* causation could not be inferred due to a small sample size. Similar to *Chapter III,* we again draw on the suggestions made in *Chapter II* by collecting rich behavioural data to explore how coordinating behaviours are impacted by familiarity amongst team members.

Abstract

Multi-team systems (MTS) involve two or more component teams working in parallel to achieve separate but related objectives in the context of collective over-arching goals. They are found in dynamic and challenging environments, where multiple specialised and highly skilled teams come together for a short period of time. Evidence from conventional teams shows that familiarity between members can enhance performance. We tested if the same holds true for MTS, and if such a relationship was due to the effects of familiarity on shared goals and communication. Data were collected from participants engaged in an immersive simulation of a terrorist incident (*N* = 22 teams). Results showed that familiar teams engaged in more positive team performance behaviours than unfamiliar teams. This effect was due, in part, to familiar teams engaging in significantly more communication, rather than through the development of shared goals. Findings are discussed with reference to the current literature on MTS and possible implications for multi-agency teams.

Keywords: *Multi-team systems, Familiarity, Communication, Shared Goals*

Multi-team Systems in Crises: Does Familiarity Matter?

Multi-team systems (MTS) comprise specialised component teams who work to accomplish separate goals in the context of collective over-arching shared goals (DeChurch & Mathieu, 2009; Mathieu, Marks & Zaccaro, 2001). MTS are formed in extreme and challenging environments in which errors in teamwork can have catastrophic consequences (Bell, Fisher, Brown & Mann, 2018; Mathieu et al., 2001). Examples of MTS include emergency response teams (Waring, Alison, Barrett-Pink, Humann, Swan & Zilinsky, 2018), disaster response teams (Rico, Hinsz, Burke & Salas, 2017), military teams (Wijnmaalen, Voordijk, Rietjens & Dewulf, 2019) and clinical care teams (Gerber et al., 2016). Interest in MTS is increasing as researchers acknowledge that findings from conventional teams do not always transfer to MTS (Luciano, DeChurch & Mathieu, 2018). Over the last decade we have seen an expansion in theoretical models and frameworks seeking to explain how multiple teams can work together within a tightly coupled system (Luciano et al., 2018; Zaccaro, Marks, & DeChurch, 2012). However, there is still much to be learned about these systems and empirical evidence in this area lags behind theoretical developments (Shuffler, Jiménez-Rodríguez & Kramer, 2015). Research is needed to begin addressing the challenges of working in multi-agency teams, and to develop practical solutions to improve teamwork in this context (Waring, Alison, Shortland & Humann, 2019; Wijnmaalen et al., 2019)

One area that has gained interest is the effects of familiarity on team processes. Research from conventional teams has shown that familiar teams tend to outperform unfamiliar teams (Espinosa, Slaughter, Kraut & Herbsleb, 2007; Huckman, Staats & Upton, 2009). This effect has been attributed, in part, to the fact that familiarity allows a team to devote their immediate focus to task demands in the early stages of team development, rather than establishing interpersonal relationships and norms (Adams & Anantatmula, 2010; Harrison, Mohammed, McGrath, Florey & Vanderstoep, 2003; Killumets, D'Innocenzo, Maynard & Mathieu, 2015). When applied to MTS, this suggests that the lack of familiarity inherent among members may compromise their success (Shuffler et al., 2015). This is because unfamiliar teams have a limited understanding of one another's roles and responsibilities, which delays the development of communication channels and reduces opportunities to coordinate on shared goals (Power & Alison, 2017; Waring et al., 2019). The current Study sought to test the role of familiarity on team behaviours in a MTS context. In addition, we examined two processes through which the effect of familiarity may occur: goal setting and communication (Figure 1). Shared goals and communication have been identified as important factors in MTS performance (DeChurch & Marks, 2006; Fodor & Flestea, 2016) and both have been linked to familiarity in conventional team research (Marlow, Lacerenza, Paoletti, Burke & Salas, 2018; Mowday, Steers & Porter, 1979).

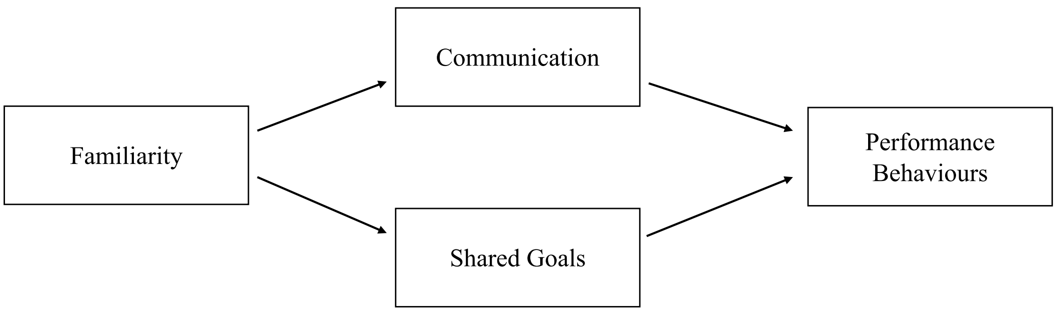


Figure 1. Hypothesised Model

Familiarity and MTS

MTS face unique challenges when completing tasks because they require the coordinated efforts of previously unacquainted teams who have little to no familiarity (Shuffler et al., 2015). Familiarity is defined as team members' shared experience of working together (Huckman, Staats & Upton, 2012) and is shown to improve team members' ability to coordinate and communicate during tasks (Espinosa et al., 2007). As team members become familiar over time, their knowledge of who knows what within the team (i.e., transactive memory system, [Austin, 2003]) becomes stronger and a greater awareness of team functioning and expected behaviours is generated (i.e., their team mental model, [Salas, Sims & Burke, 2005]). Both transactive memory systems and team mental models improve coordination mechanisms within the team (Espinosa et al., 2007; Okhuysen 2001). In contrast, a lack of familiarity is associated with an increase in accidents (National Transportation Safety Board, 1994) and a reduced willingness to accept assistance from team members (Smith-Jentsch, Kraiger, Cannon-Bowers & Salas 2009).

The benefits of familiarity have been noted across a range of contexts including sport (Dalal, Nolan & Gannon, 2017; Sieweke & Zhao, 2015), medical care (Joshi, Hernandez, Martinez, Abdel-Fattah & Gardner, 2017; Reagans, Argote & Brooks, 2005) software development (Huckman, Staats & Upton, 2009) and virtual teams (Maynard, Mathieu, Gilson, Sanchez & Dean, 2019). There is indirect evidence that team member familiarity may also benefit MTS. For example, inter-agency training, which brings together component teams to improve cross-team working and generate familiarity of how different agencies operate, is proposed to alleviate the challenges of working in MTS (Gerber et al., 2016; Waring et al., 2019). Without inter-agency training, teams often need more time to establish appropriate communication channels and identify which team members hold the appropriate knowledge for task completion. These types of "orientation" tasks can negatively impact performance as they often lead to delays in decision-making (Luciano et al., 2018; Shuffler et al, 2015). Cujipers et al. (2016) also found that when team members spent more time with one another, they identified more strongly with the MTS, which increased subsequent performance. Taken together, we predict that:

Hypothesis 1. Familiar teams will show significantly more positive team performance behaviours and significantly fewer negative team performance behaviours than unfamiliar teams.

We focus on comparing team behaviours in familiar and unfamiliar teams (specifically, actions that are relevant to task completion and reflect interactions between team members), rather than task outcomes (consequences of performance behaviours). This is because the latter can be difficult to measure in extreme environments where decisions may have no right or wrong answer and outcomes can be dictated by external factors in the environment (e.g., number of victims involved in incident, a secondary explosive device) (Alison, Power, van den Heuvel & Waring, 2015a). Measuring performance by team behaviours (e.g., number of joint decisions) provides richer data when compared to objective measures (e.g., what decision was made), which may be deficient in providing a comprehensive metric of team dynamics (Salas, Reyes & McDaniel, 2018).

Shared Goals, Communication and MTS

Research suggests familiarity may impact performance indirectly through team processes (Killumets et al., 2015). Two processes identified by previous research are shared goals (Mowday et al., 1979) and communication (Maynard et al., 2019). Shared goals are defined as conscious purposes or deliberate plans (Locke & Latham, 2002). They facilitate effective team performance when they are specified at the outset and accompanied by sub-goals needed for task completion (Mathieu et al., 2001). In MTS, entirely shared goals are unlikely as each team pursues specific intra-team goals in conjunction with superordinate, system-level goals (Luciano et al., 2018). However, a critical feature of the MTS is the ability of teams to manage this, ensuring shared system-level goals can be achieved alongside component team-level goals (Bienfeld & Grote, 2014; Shuffler et al., 2015).

System-level goals and component team goals structure for MTS as a "hierarchy of goals": essentially, a network of interrelated goals, including both distal (system level) and proximal (component team) goals (Mathieu et al, 2001). When goals are incompatible and dissimilar across component teams, MTS are said to have high levels of *goal discordancy,* which is disruptive to effective teamworking (Luciano et al., 2018). This has been shown in studies where MTS performance improved when leaders were able to synchronise goals across component teams (DeChurch & Marks, 2006). Further, it is supported by studies with conventional teams where shared goals are associated with increased performance (Bradley, Foltz, White & Wise, 2006; Colbert, Kristof-Brown, Bradley & Barrick, 2008; Gully, Devine & Whitney, 1995).

Research in different areas suggests that shared team goals may mediate the effect of familiarity on team behaviours in MTS. Research shows that working with other team members increases commitment towards team values and shared goals (Mowday et al., 1979). Relatedly, research on team mental models suggests that as team members become more familiar with one another, they can perform their individual activities in a way that is consistent with team goals and the needs of other team members (Weick & Roberts, 1993). Further, familiarity is known to increase the strength of social identity in organisational teams (Chang, Chang & Sha, 2009), with identification leading individuals to favour its members and crucially its goals over and above other groups (Tajfel, 1982). In MTS, members of component teams often identify more strongly with their own team in place of the wider system and this can reduce their willingness to pursue shared superordinate goals (Connaughton, Williams & Shuffler, 2012). Empirical evidence suggests the more team members identify with the MTS, the more likely they will be to invest in achieving shared superordinate goals (Cuijpers et al., 2016). Cuijpers et al. (2016) propose that one way to improve the extent to which team members identify with the MTS is to increase familiarity across the system.

Hypothesis 2. Familiarity in teams will indirectly impact team performance behaviours by increasing commitment to shared goals.

Communication is a key part of an effective MTS as it enables members to share information and develop shared meaning (Keyton, Ford & Smith, 2012). Effective communication facilitates collaboration and the achievement of superordinate goals (Davison, Hollenbeck, Barnes, Sleesman & Ilgen, 2012; Keyton et al., 2012), enables teams to develop shared situation awareness (Salas, Prince, Baker & Shresthsa, 1995) and improves coordination (Firth, Hollenbeck, Miles, Ilgen & Barnes, 2015). Of particular interest is the structure and distribution of communications in MTS. Research has shown that dense communication networks lead to high goal attainment and performance (Balkundi & Harrison, 2006). Evidence also suggests that decentralised MTS networks (i.e., networks in which communications are evenly distributed across team members) may be beneficial in crises (Brown & Power, 2018; Lanaj, Hollenbeck, Ilgen, Barnes, & Harmon, 2013). By adopting a decentralised structure, information is shared widely across the system making it more accessible to different team members (Schraagen & Van de Ven, 2011).

The relationship of open and decentralised communication networks with effective performance in MTS has been established in previous studies (Keyton et al., 2012; Lanaj et al., 2013). There is further evidence that the relationship between communication and performance is related to familiarity. Maynard et al. (2019) identified communication as the mediating mechanism linking familiarity to team performance in a study of 63 distributed virtual teams. Their research showed that team member familiarity increased information elaboration, which increased team effectiveness as rated by both managers and team-leaders. The positive effect of communication on the familiarity-performance relationship has also been demonstrated in a meta-analysis of 95 studies, in which it was found that the relationship between communication and performance strengthens as familiarity increases (Marlow et al., 2018). Further, an absence of familiarity has been shown to hamper the amount of information shared within teams (Gibson & Cohen, 2003).

Familiar teams spend more time sharing task-related information as they require less time to acquire inter-personal information (Harrison et al., 2003). This is advantageous for MTS responding in crises as it enables them to quickly deal with the task at hand. Familiarity can also improve communication between team members due to better developed transactive memory systems (Ren & Argote, 2011). For MTS, a greater understanding of each other's roles and capabilities during crises may lead to a timely and appropriate exchange of information (Waring et al., 2018). Further, familiarity is shown to influence network structures in teams, such that pre-existing relationships can drive the way in which team members interact with one another (i.e., density of interactions and network centralisation) (Grund, 2016; Soda, Usai & Zaheer, 2004).

Hypothesis 3. Familiarity in teams will indirectly impact team performance behaviours by increasing the effectiveness of communication networks, as indicated by frequent communication, denser networks and less centralised networks.

Individual differences

In addition to structural features (e.g., familiarity), research suggests teamwork may be influenced by other aspects of team composition such as the individual differences of team members (Bell, 2007; Bell, Brown, Colaneri & Outland, 2018; Glew, 2009). Teams comprise different individuals, thus the skills, knowledge and characteristics of each team member will have important implications for how teams interact (Hackman, 1987). Of particular interest are personality traits and values (Bell, 2007; Vessey & Landon, 2017). Personality, as measured through trait elevation (i.e., aggregated personality traits within the team [Kozlowski & Klein, 2000]), has been show to impact team performance when a team is high in extraversion and emotional stability (Bell, 2007; Barrick, Stewart, Neubert & Mount, 1998). However, these findings are not universal, and some meta-analyses have reported non-significant relationships (Peeters, Van Tuijl, Rutte & Reyman, 2006). Values are defined as enduring generalized beliefs which guide individuals' behaviour (Kabanoff, Waldersee & Cohen, 1995; Meglino & Ravlin, 1998). Values that are socially oriented (e.g., high tolerance of others), increase cooperation and performance in teams (Eby & Dobbins, 1997; Kirkman & Shapiro, 2001). In the current Study, we examined the effects of trait elevation in extraversion, openness to new experiences, conscientiousness, agreeableness and emotional stability, in addition to socially oriented and self-oriented values. This was to test if individual differences were important to performance in MTS, but also to establish if familiarity predicted team behaviours after these individual differences were controlled.

Method

Participants

One hundred and one undergraduate students (71 female, 30 male) were recruited though opportunity sampling and by targeted sampling from sports societies. The Study was advertised as a simulation on decision-making during emergencies, offering participants £7 for taking part. The average age of participants was 21 years (*SD* = 2.90; Range 18 – 38). Participants were placed into 22 teams; nine teams comprised four members, and 13 teams comprised five members. Of the 22 teams, 11 teams were “familiar” (sports teams who met a minimum of twice a week and had been together as a team for at least four months). The remaining 11 “unfamiliar” teams comprised members taken from the general student population. To ensure participants in the unfamiliar teams did not know each other, they were asked to sign up for multiple study slots, and to state their University course and year of study. Those who stated the same discipline in the same year were allocated to different groups. When participants arrived to take part, a final verbal check was made to ensure that participants in the unfamiliar condition did not know one another, and that participants in the familiar condition were all part of the same sports team/society. This Study was approved by the Faculty of Science and Technology Research Ethics Committee at Lancaster University (FST18007).

Measures

Personality. Personality traits were measured with the Ten Item Personality Scale (Gosling, Rentfrow & Swann, 2003). The scale measures the Big-Five personality traits: Extraversion (α = .72), Openness to new experiences (α = .48), Conscientiousness (α = .70), Agreeableness (α = .34) and Emotional stability (α = .61)[3]. Participants are presented with two traits (e.g., Extraverted/Enthusiastic) and asked to rate the extent to which the *pair* of traits applies to them. Responses are made on a seven-point likert scale from one (strongly disagree) to seven (stronlgly agree). The five traits were aggregated to the team level by generating a mean score for each trait per team.

Values. Values were measured using the Portrait Value Questionnaire (Schwartz, 2006). The measure comprises 40 items that measure ten values: Conformity, Tradition, Benevolence, Universalism, Self-Direction, Stimulation, Hedonism, Achievement, Power, and Security. Participants state the extent to which they feel a description applies to themselves, on a six-point likert scale from one (not like me at all) to six (very much like me). Example questions include: *"They like suprises. It is important to them to have an excting life"* (Stimulation) and *"It is important for them to be independent. They like to rely on themselves"* (Self-Direction). Cosistent with others (Hinz, Brähler, Schmidt & Albani, 2005; Schwartz, 2011), we grouped these values into those that reflect "social orientation" (Conformity, Tradition, Benevolence, Universalism, Security; α = .64) and those that refelct "self-orientation" (Power, Achievement, Hedonism, Stimulation, Self-direction; α = .48). An inspection of the self-orientation subscale item correlations suggests that inernal consistency may be improved by removing Power values. As such, Power was removed before

[3] The low internal consistency estimates are consistent with other reported uses of this scale and are attributed to the fact that the scale measures broad domains with only 10 items. It is a measure designed for instances where time is short and personality is not the sole topic of interest (Gosling et al., 2003) (https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/a-note-on-alpha-reliability-and-factor-structure-in-the-tipi/)

producing a composite score for the self-orientation scale (α = .54). Collated scores for social orientated values and self-oriented values were aggregated to the team level by generating a mean score for each sub-scale per team.

Shared goals. Participants were required to state their individual goal priorities on a paper-based decision log at two set time points in the simulation; after Inject 1 (Time 1) and Inject 8 (Time 2) (*see* Table 1, Figure 2). Participants were asked to consider the information that had been provided to them and to each write down their five main individual goal priorities at that time. Participants were not given a prescriptive list of goal priorities and were not instructed to write them in order of importance. The goals recorded at Time 1 were used as a control measure, to ensure no significant differences between familiar and unfamiliar teams at the start of the experiment (t (20) = -.60, p = .40).

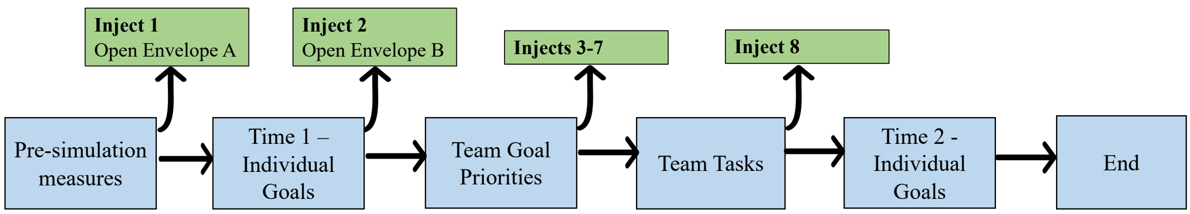


Figure 2. Study Procedure

To obtain a measure of shared goals (a mediator in the current Study), we counted the number of individual goals recorded at Time 2 that were congruent with the five Team Goal priorities each group provided earlier in the experiment (following Inject 2). For example, if a Team Goal priority was to evacuate and treat casualties and a participant included treating casualties in their individual Goals at Time 2, this would count as a shared goal. The mean number of shared goals per group were used in the analysis as a mediator. This approach to measuring shared goals is similar to Bradley et al.'s (2006) Goal Congruence measure. To ensure individual goals were reliably coded as being congruent with Team Goal Priorities, a second person coded 20% of the data, with high inter-coder reliability (k = .81).

Table 1. Simulation injects

Inject	Format	Description
1	Phone call	Open Envelope A. Participants receive a phone call from the emergency services control centre outlining that there has been an explosion at the finish line of the marathon and there are approx. 2000 individuals in the immediate area. Participants are asked to state their five main goal priorities individually (e.g., save life, evacuate surrounding area).
2	Written message	The researchers ask participants to introduce themselves (in their role) to the other team members and to decide, as a team, what their five main shared goals priorities are.
3	Radio	Open Envelope B. Operational police commander on scene sends a radio message requesting assistance zoning the incident ground.
4	Phone call	Participants receive a phone call from the Strategic lead of the incident instructing them to select an appropriate media message to send out to the public.
5	Twitter feed	Tweets criticising the response of the emergency services. Designed to increase the sense of urgency.
6	Radio	Participants receive a phone call from Ambulance commander. Non-specialist responders are treating casualties in a high-risk zone. Participants must decide whether to follow procedure and pull back the responders.
7	Phone call	The Council Emergency Planning committee call to request that participants identify an appropriate reception for those involved in the incident and their family members.
8	Phone call	Participants receive a second phone call from the Strategic lead asking them to state, individually, their five main goals priorities

Communication networks. To measure communication, audio data from each of the simulations were transcribed verbatim and coded into communication matrices. Matrices indicated which team member spoke to whom and with what frequency. The matrices were loaded into R (R Core Team, 2012) and density estimates were calculated using the "sna" package (Butts, 2008). As networks were weighted (i.e., they measured the *number* of interactions between team members), network centralisation was calculated using the eigen vector centrality measure in UCINET (Bihari & Pandia, 2015; Borgatti, Everett & Freeman, 2002; Wei, Pfeffer, Reminga & Carley, 2011).

Centralisaton, density and the sum of communications between team members were used as mediators.

Team Performance. Consistent with previous research (Waller & Kaplan, 2018; Westli, Johnsen, Eid, Rasten & Brattebø, 2010), team “performance” was measured through interaction behaviours during the simulation, rather than decisions made, as the latter were not designed in the Study to have a right or wrong answer. We used Brown and Power’s (2018) coding dictionary of coordinating behaviours, that classifies behaviours as positive (Joint Decision Making, Sharing Task Related Information and Sharing Resources) or negative (Decision Uncertainties, Role Uncertainties and Conflicting Priorities) (*see* Table 2). The codes were identified using deductive thematic analysis in Nvivo (*see* Appendix 3, Braun & Clarke, 2006; Nowell, Morris, White & Moules, 2017). After all transcripts had been coded by the first author, they were re-coded to check for intra-rater agreement (k = .75). Four (18%) transcripts were also coded by an independent researcher blind to the group (familiar/unfamiliar) that the data were taken from. The coder was trained prior to coding and the results showed a good level of inter-rater reliability (k = .72). The six team behaviours are considered as independent outcome measures in the analyses.

Table 2. Positive and Negative MTS coordinating behaviours (Brown & Power, 2018).

Performance behaviours	Description
Joint decision-making (JDM)	Instances in which the team actively worked together to implement a decision
Sharing resources (SR)	Instances in which team members offered resources to assist other agencies.
Sharing task related information (STRI)	Instances in which team members actively sought to share role specific information and improve shared situational awareness.
Role uncertainties (RU)	Confusion about one's own role and the role of others.
Decision uncertainties (DU)	Instances in which misinformation led to confusion and discussions were delayed due to indecision.
Conflicting priorities (CP)	Instances in which team members attempted to re-orient the conversation away from inter-team priorities and towards intra-agency priorities.

Procedure

On arrival at the lab, participants sat around a 'group' table in the centre of the room. They listened to a short presentation informing them they would be taking part in a simulated major incident and given some examples of what this might be (e.g., a major traffic accident, a terrorist incident). They were briefed about the structure of incident response in the U.K. and informed they would be role playing at the tactical level of response. Tactical responders typically meet at a location close to the incident site, ensuring operational commanders on the ground conduct a coordinated cohesive response, whilst liaising with strategic command to ensure the overall strategy is followed. Participants were instructed that, consistent with tactical level response, they would be responding to the incident as a MTS in which each team member would be assigned a different agency role (e.g., Police, Fire). Participants were informed that during the simulation, all information, instructions, audio clips and images would be presented to them on a television in the corner of the room and that the researchers would be in a control room next door observing via video cameras situated throughout the room. At this point, participants were invited to ask any questions and invited to withdraw from the Study if they did not feel comfortable. All participants chose to remain and gave their constent to continue. Following this, participants were asked to complete an online pre-simulation questionnaire, using tablets provided by the researchers. On completion of the questionnaire, participants were given a paper "individual decision log" which they were told to complete in silence during the simulation when instructed to do so. In addition, participants were shown an "electronic team decision log" which was a computer stationed on a table at the side of the room, that the group used to record group decisions with the researchers over Skype messanger and to ask questions should they need to. The researchers then left the room.

The scenario was based on the response to a terrorist attack at the finish line of a marathon in a city in the North of England. The simulation was developed by the researchers, with input from commanders from the Police, Fire and Ambulance service. It was designed to be suitable for students, whilst reflecting the challenges of a major incident. At the start of the simulation, team members were instructed (via the television) to take an envelope on the table at random and to remain in silence until further instruction (*see* Figure 2). The envelopes were labelled "A" and numbered 1-5 (each number denoted a different role in the simulation: Police, Fire, Ambulance, Mayor or Marathon Organiser), ensuring random assignment of participants to roles. As shown in Table 3, the envelope contained information about their role and typical agency responsibilities (e.g., the Police are responsible for preserving the scene and speaking to witnesses). In instances where teams comprised four members, roles of Mayor and Marathon Organiser were combined and participants were instructed to act as if the Mayor had been involved in the planning of the marathon. This was done for two reasons. First, combining the other roles (Fire, Police and Ambulance) lacks ecological validity whereas combining Mayor and Marathan organiser is more feasible.

It is more likely for a Mayor to have had oversight in the council planning and preparations of a Marathon than for the blue-light services to be involved in this. Second, leaving a role out would result in some information not being accessible to the team, thus making teams unequally matched when making decisions. In total, nine teams were requried to combine roles (six = Unfamiliar, three = Familiar). Mann-Whitney U tests were carried out to compare the effects of team size on the behaviours in familiar teams and unfamiliar teams repsectively. No significant differences were identified in any of the performance behaivours in familiar teams of four or five (avg. $U = 6.50, p = >.05$) and in unfamiliar teams of four and five (avg. $U = 11.42, p = >.05$)

At Inject 1, and after reading information on their role and typical responsibilities using the information in Envelope A (*see* Table 3), participants were asked to make a written list of their individual five main goal priorities without speaking to one another. After five minutes, participants were instructed to introduce themselves (in their role) and decide, as a team, what their collective five main goal priorities were (Inject 2). Once the teams had reached consensus on five team goal priorities, they were each provided with additional role specific infromation in Envelope B (*see* Table 3) to inform their decsion-making for the remainder of the simulation (e.g., Police were informed that they have 4 Armed Response Vehicles in the immediate area). This information was not provided at the outset, so as not to prime participants when deciding their individual and team goals (e.g., for the role of Mayor, participants were informed in Envelope B that they had the responsibility to take care of local businesses within the vicinity of the incident).

Table 3. Overview of information provided in Envelope A and B

Role	Envelope A	Envelope B
Police	Overview of the role of tactical command. Outline of Police specific priorities: neutralising threats to life, road closures, speaking to witnesses, crime scene preservation etc.	Information on resources such as the number of Armed Police in the vicinity, instructions on how to manage communications on causalty numbers and site maps.
Fire	Overview of the role of tactical command. Outline of Fire specific priorirites: extracting casualties, ground safety etc.	Information on resources such as the number of specialist trained Search and Rescue responders.
Ambulance	Overview of the role of tactical command. Outline of Ambulance specific priorities: prioritisation of causalties, liaising with NHS England on hospital capacity etc.	Information on resources such as the number of specialist trained Hazardous Area Response Team members. Capacity of nearby hospitals and maps identifying where the hospitals are located.
Mayor	Overview of the role of tactical command. Outline of Mayor specific priorities: appropriate shelter for those involved in the incident, liaise with the coroners office to establish an emergency mortuary etc.	Information on voluntary agencies (e.g., the Red Cross) and local businesses in the area that may be affected. Maps detailing the location of Reception Sites.
Marathon Organiser	Overview of the role of tactical command. Outline of Marathon organiser specific priorities: the wellbeing and safety of staff members on site, the legacy of the event, manging incoming communications with sponsors etc.	Information on the number of staff members present on site of explosion and the number of marathon runners. Maps of the marathon course and detailed maps of incident site.

The simulation continued with audio, written and visual injects (as per Table 1 and Figure 2), with teams given key tasks at Inject 3, 4, 6 and 7. For each of these tasks, participants were given a series of options. Each task was designed to be challenging and provoke discussion, with teams guided by the information provided in Envelope A and B (Table 3). A maximum time of 7 minutes was given for each of the key tasks. The decision to give 7 minutes per task was based on the time taken during a pilot simulation and discussion with responders from the emergency services. Teams were invited to make additional decisions as they saw fit (e.g., requesting additional armed Police officers on scene). Teams continued to discuss and complete tasks until Inject 8. At this point, participants received a phone call from the Strategic lead (one of the researchers) asking them to reflect on the incident as a whole and in silence note down their five main goal priorities for their role. They were told this would signal the end of the simulation. Simulations lasted between 39 minutes and 53 minutes (*M* = 46.14*, S.D* = 4.19*)*, dependent on the level of discussion in the team and were not significantly different for familiar (*M* = 47.27*, S.D* = 3.85*)* and unfamiliar (*M* = 45.82*, S.D* = 4.73*)* teams, (*t* (20) = .79, *p* = .49). Each simulation was audio recorded. Following the simulation participants were de-briefed by the research team.

Results

The main aim of the Study was to establish if familiar MTS demonstrated significantly more positive and significantly fewer negative team performance behaviours than unfamiliar MTS. Table 4 shows the relationships between these variables. Consistent with Hypothesis 1, familiar teams engaged in more positive team behaviours relating to *joint decision-making*, *sharing resources,* and *sharing task related information*. However, and in contrast to our prediction, they also displayed more instances of *conflicting priorities*. In addition to familiarity, we found that teams who shared more resources and task related information were high in extraversion; whilst those sharing less resources were high in agreeableness and/or high in social orientated values.

To test if familiarity was a significant predictor of sharing resources and task related information after controlling for the trait and value composition of the group, we carried out two hierarchical regression models. The personality and value variables that were significantly related to the performance behaviours were entered in Block 1, and familiarity was entered in Block 2. The results of the regression analyses, as reported in Table 5, show that familiarity predicts *sharing task related information* after extraversion is controlled. However, it does not predict *sharing resources* when a team's agreeableness, extraversion and socially oriented value composition is controlled for. In the case of *sharing resources*, all the variables emerge as non-significant, and whether a team is familiar no longer influences these behaviours.

Hypotheses 2 and 3 predict that familiarity is indirectly related to performance behaviours through increased shared goals and communication (increased density, higher sum of communication and less centralised networks). To test these predictions, we focused on *joint decision-making*, *sharing task related information* and *conflicting priorities*, as these behaviours were found to be uniquely related to familiarity. Figures 3-5 show the relationships between familiarity, shared goals, communication and the team behaviours. As can be seen, shared goals are not related to familiarity or any of the three team behaviours. Hypothesis 2, that shared goals mediate the effects of familiarity, is therefore not supported.

Table 4. Means, SD, and correlations between all Study variables (*N*=22 teams)

	Familiar		*Unfamiliar*																			
	M	*SD*	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Familiarity	N/A	N/A			-																	
2. Extraversion	4.92	.80	4.19	.66	.46*	-																
3.Emotional Stability	4.41	.63	4.29	.33	.16	.37	-															
4. Conscientiousness	4.77	.81	5.33	.67	-.38	.03	-.20	-														
5. Openness	5.03	.41	5.07	.48	-.05	.27	.09	.09	-													
6. Agreeableness	4.44	.47	4.68	.52	.25	-.26	.08	.02	.27	-												
7. Social orientation	19.34	1.31	20.62	.76	.53*	.03	.03	.28	.02	.36	-											
8. Self-orientation	17.69	1.30	17.61	1.81	.33	.08	.08	.09	.40	.14	.23	-										
9. Shared Goals	2.17	.74	2.02	.56	.12	-.28	-.26	-.01	-.36	-.02	-.10	-.04	-									
10. Sum of Comms	582.27	123.96	264.36	67.86	.86*	.44*	.05	-.25	.02	-.25	-.54*	.07	.19	-								
11. Density	33.22	5.57	17.38	5.57	.83*	.51*	-.01	-.15	.01	-.15	-.62*	.16	.07	.83*	-							
12. Centralisation	26.82	6.89	35.02	13.03	-.38	-.31	-.06	.36	-.05	.36	.24	-.35	.22	-.25	-.50*	-						
13. Joint decision making	14.46	3.96	9.10	2.98	.63**	.09	.09	.12	.07	-.35	-.21	.25	.28	.67*	.49*	.13	-					
14. Sharing resources	2.55	1.70	.82	1.08	.54**	.53*	.40	.06	-.08	-.46*	-.51*	-.05	.01	.65*	.60*	-.13	.56*	-				
15. Sharing task related information	18.36	4.10	12.81	1.08	.62**	.54*	.23	-.10	-.03	-.41	-.36	.06	-.01	.72*	.60*	-.09	.52*	.64*	-			
16. Role uncertainties	1.55	1.81	.82	1.25	.24	-.09	-.21	-.30	.05	-.01	-.32	.11	.33	.38	.17	-.10	.12	.10	-.01	-		
17. Decision uncertainties	2.91	2.59	2.36	2.66	.11	-.46*	.02	-.42	.13	.13	-.24	-.34	.01	.03	.12	-.13	-.22	-.14	-.18	.28	-	
18. Conflicting priorities	3.36	2.25	1.36	1.12	.51*	.32	-.12	-.27	.22	-.30	-.23	.30	.11	.73*	.50*	-.32	.59*	.31	.52*	.40	-.05	-

Note: *significance = <.05, ** = significance <.001

T**able 5.** Hierarchical regression analysis results for familiarity and team behaviours

	R^2	ΔR^2	β	p
Sharing task related information	.454	.258		
Block 1				
Extraversion			.542	.009
Block 2	.687	.416		
Extraversion			.327	.097
Familiarity			.474	.011*
Sharing resources	.705	.413		
Block 1				
Agreeableness			-.232	.221
Socially oriented values			-.346	.072
Extraversion			.404	.032
Block 2	.720	.406		
Agreeableness			-.235	.220
Socially oriented values			-.257	.233
Extraversion			.333	.102
Familiarity			.192	.196

Note: significance = <.05*, significance = <.01**, (one-tailed for the effect of Familiarity), N = 22.

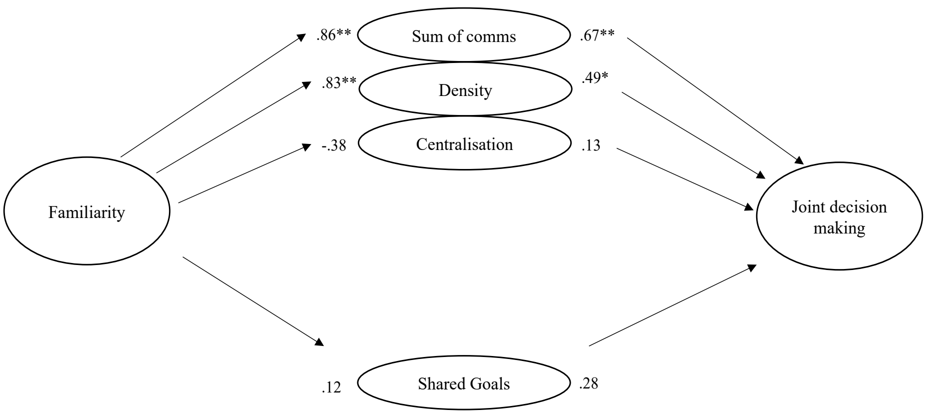


Figure 3. Correlations in hypothesised model for Joint Decision-Making

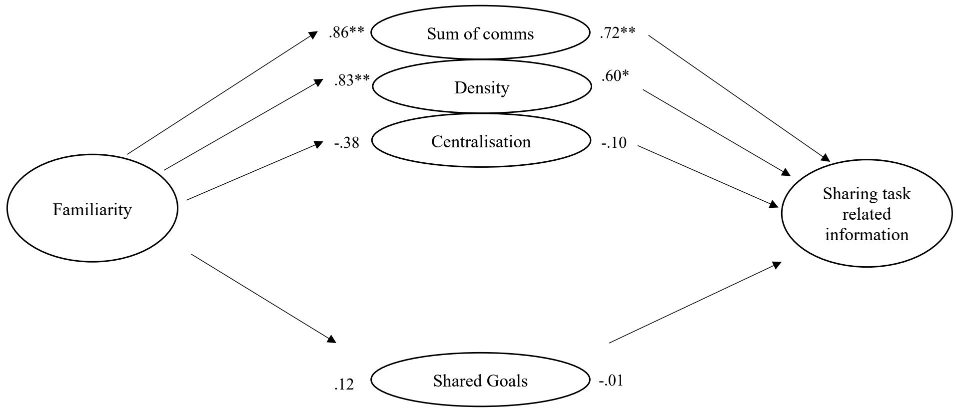


Figure 4. Correlations in hypothesised model for Sharing task related information

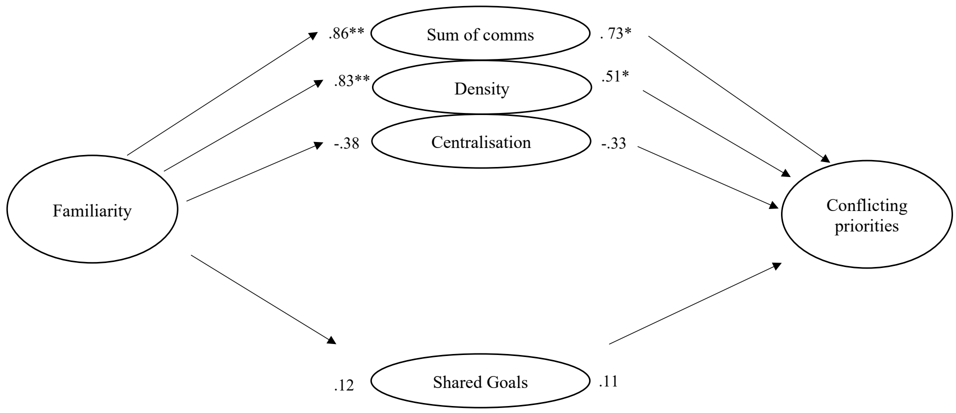


Figure 5. Correlations for hypothesised model for Conflicting Priorities

To test the role of communication as a mediator of the effects of familiarity on performance, we examined the sum of communications, and the density and centralisation of the communication networks. Of these three measures, the results show significant relationships for sum of communications and density with familiarity and the three performance behaviours. To test the indirect effect of familiarity on team behaviours through sum of communications and density, we used Schoemann, Boulton and Short's (2017) approach for team data. As per this approach, we first ran a Monte Carlo Power Analysis on the independent variable, mediators and dependent variables to see if there was sufficient power to test for indirect effects. Sufficient power was obtained for models testing the indirect effect of familiarity through sum of communications on *sharing task related information* (β = .71) and *conflicting priorities* (β = .94) only. For both of these models, indirect effects analysis was conducted using the R package "lavaan" (Rosseel, 2012). A bias correct bootstrapped interval with 10,000 samples was used to estimate 95% confidence intervals.

The results showed that familiarity had an indirect effect on *sharing task related information* through sum of communications. Specifically, familiarity was associated with an $a = 317.91$ ($SE = 40.63$), $p < .001$, increase in sum of communications. Controlling for familiarity, an increase in sum of communications was associated with an increase in *sharing task related information*, $b = .019$ ($S.E = .007$), $p = .006$. For every $a = 317.91$ in the association between familiarity and sum of communications, there was an $ab = 5.82$ ($SE = 2.72$) increase in the amount of *sharing task related information*, $p = .010$, CI [1.43, 10.00]. The direct relationship between familiarity and *sharing task related information* was non-significant, $c = -.37$ ($SE = 2.49$), $p = .89$.

The results showed that familiarity had an indirect effect on *conflicting priorities* through sum of communications. Specifically, familiarity was associated with an $a = 317.91$ ($SE = 40.63$), $p < .001$, increase in sum of communications. Controlling for familiarity, an increase in sum of communications was associated with an increase in *conflicting priorities*, $b = .012$ ($SE = .003$), $p = <.001$. For every $a = 317.91$ in the association between familiarity and sum of communication, there was an $ab = 3.71$ ($SE = 1.03$) increase in *conflicting priorities*, $p = .010$, CI [1.70, 5.72]. The direct relationship between familiarity and *conflicting priorities* was non-significant, $c = -1.71$ ($SE = 1.06$), $p = .11$. Both models support Hypothesis 3.

For those models where there was insufficient power to test for indirect effects (i.e., familiarity through density on all three performance behaviours, and familiarity through sum of communications on *joint decision-making*), we ran four hierarchical regression analyses. Familiarity was entered in Block 1, and the relevant mediator(s) was entered in Block 2. As shown in Table 6, sum of communications and density were non-significant predictors of all three performance behaviours when entered in a model with familiarity. This suggests that they are unlikely to offer a route through which familiarity predicts these behaviours.

Table 6. Hierarchical regression results for the relationship between familiarity, communication measures and team behaviours

	R^2	ΔR^2	*b*	*p*
Joint decision-making				
Block 1	.392	.361		
Familiarity			.626	.002*
Block 2	.457	.400		
Familiarity			.200	.551
Sum of communications			.497	.147
Joint decision-making				
Block 1	.392	.361		
Familiarity			.626	.002*
Block 2	.395	.331		
Familiarity			.705	.040*
Density			-.095	.769
Sharing task related information				
Block 1	.387	.356		
Familiarity			.622	.002*
Block 2	.410	.348		
Familiarity			.394	.228
Density			.274	.397
Conflicting priorities				
Block 1	.508	.221		
Familiarity			.508	.016*
Block 2	.529	.204		
Familiarity			.289	.418
Density			.264	.459

Discussion

Operating in MTS can be challenging as previously unacquainted teams are brought together swiftly to complete complex tasks (Shuffler et al., 2015). The current Study examined if familiarity can help in this process. Further, it examined if the effects of familiarity occur indirectly through increases in shared goals and effective communication channels. The results showed that familiar teams engaged in more *joint decision-making*, *sharing resources* and *sharing task related information*. Findings provide the first empirical validation of the benefits to familiarity in MTS and support suggestions that increasing inter-agency training will be of use to MTS who operate in crises (Gerber et al., 2016; Waring et al., 2018; Wijnmaalen et al. 2019).

In addition to positive behaviours, we found that familiar teams also expressed more conflicting priorities. Conflicting priorities are identified as detrimental to performance (Hinsz & Betts, 2012), yet there was no indication that familiar teams in the current Study were performing worse than unfamiliar groups (familiar teams had higher instances of all three positive performance behaviours). This suggests that *conflicting priorities* may have had a positive effect on team interactions; a finding that is in line with some research on voicing behaviours. Voicing out and *conflicting priorities* are both associated with instances in which team members speak out and voice their opinion. Voicing out has been shown in other studies to be conducive to performance as it is indicative of teams sharing a sense of psychological safety; a confidence to share critical information in the knowledge it will not be used against them (Edmondson, 1999; Staats, Gino & Pisano, Edmondson, Pierce & Spektr, 2010). Familiarity is a key driver in the development of psychological safety suggesting the impact *of conflicting priorities* on team efficacy may be moderated by psychological safety (Roberto, 2002). Specifically, when teams share a sense of psychological safety, *conflicting priorities* can play a constructive role, reflecting critical discussion and confidence to challenge one another's ideas.

Contrary to our expectations, no effect of familiarity was identified for *decision uncertainties* and *role uncertainties*. This is surprising, as familiar teams possess a greater knowledge of one another's working (e.g., less role uncertainty) and will anticipate one another's behaviours better (e.g., less decision uncertainty) (Austin, 2003; Salas, Sims & Burke, 2005). This finding may be attributed to an absence of task familiarity (prior experience working on a specific task) in our non-expert sample (Littlepage, Robison & Reddington, 1997). It is likely any benefits of familiarity in reducing uncertainties will have been masked by the fact all team members were enacting novel roles (as emergency responders) and engaging in novel tasks relating to major incident response.

Interestingly, the results suggest the effects of familiarity are not equal across all familiar teams. Familiarity had minimal added benefit on *sharing resources,* when accounting for the effect of individual differences. Our results showed that teams with high levels of extraversion shared more resources, while teams who were agreeable and socially oriented shared fewer resources. Extraverted team members are more likely to seek help and offer support to the team when needed (Porter, Hollenbeck, Ilgen, Ellis, West & Moon, 2003), thus they are more likely to share resources. Research also suggests socially oriented and agreeable teams should promote cooperation (such as sharing resources) (*see* Earley & Gibson, 1998; Graziano, Hair, & Finch, 1997), yet they failed to do so in this Study. One possible explanation is that agreeable and socially oriented individuals strive to remain likeable and maintain harmonious social relations, so it is possible they will have been less willing to interrupt and explicitly offer resources to others (Digman, 1990; Schwartz, 2012). Aside from *sharing resources*, the effects of individual differences on team behaviours were not strong, suggesting there are other that factors play a more important role in MTS performance.

We examined the relationship between familiarity and team behaviours through the effects of two team processes: shared goals and communication. Evidence suggests familiar teams are more committed to shared team goals than unfamiliar teams

(Cuijpers et al., 2016; Mowday et al., 1989), and so we expected to see higher instances of shared goals in familiar teams. However, we found no evidence in support of this: differences in shared goals between familiar and unfamiliar teams were marginal and both familiar and unfamiliar teams had low frequencies of shared goals. We also found no support for a relationship between shared goals and team performance behaviours, contrary to prior research (DeChurch & Marks, 2006). In DeChurch and Mark's study, shared goals were measured through the synchronization of *actions* to achieve shared goals, rather than an explicit measure of how congruent individual goals were to team goals (as in the current Study). This suggests that, whilst it is important for MTS to specify shared goals at the outset of task completion (Marks et al., 2001), as time progresses, performance behaviours are unlikely to be affected by how congruent component team goals are to shared MTS goals. It is possible this is due to entirely congruent goals preventing component teams from attending to intra-agency priorities which would detract from performance. To explore this possibility, future research may focus on the *compatibility* of component team goals to shared MTS goals (*see* Cronin & Wiengart, 2007), as this would measure how aligned teams are to the over-arching system objectives, whilst simultaneously enabling them to attend to agency specific priorities. Adjusting the framing in this way may identify differences in the goals of familiar and unfamiliar teams, as we would expect familiar teams to remain more mindful of, and therefore have goals more compatible with, system level objectives (Weick & Roberts, 1993).

The second process we considered was communication. In line with previous research, we found that familiar teams communicated more frequently, had denser communication networks and had networks that were less centralised (Gruenfeld, Mannix, Williams, & Neale, 1996; Soda et al., 2004). Our results further showed highly dense networks and networks in which communication occurred with high frequency (sum of communications) were particularly important for increasing *joint decision-making*, *sharing task related information* and *conflicting priorities*. Findings in relation to positive behaviours are in line with previous theoretical and empirical research identifying communication as a key factor driving MTS performance (Davison et al., 2012; Keyton et al., 2012; Waring et al., 2019), and shows communications are enhanced by increasing the familiarity of team members. However, we found that network centrality (the relative distribution of communications across team members) did not predict any team behaviours. In larger, disparate teams, de-centralised networks enable team members to access different streams of information (Schraagen & Van de Ven, 2011). In this Study (and for some MTS), team members were situated in the same room. Accordingly, even if individuals did not contribute to discussions equally, they remained aware of all of the information that was being shared. Findings therefore suggest when MTS operate in close quarters, centrality is unlikely to be as important as in larger, geographically dispersed teams.

Building on previous research, we predicted that communication would mediate the relationship between familiarity and team performance behaviours (Maynard et al., 2019). In support of this, we found that familiarity had an indirect impact on *sharing task related information* through the extent of communication with the team. Familiar teams communicated with a higher frequency, which resulted in the teams sharing more task-related information; allowing them to effectively pool more information (Gruenfeld et al., 1996). While sum communication and *sharing task related information* are conceptually similar, at a practical level they can be distinguished by the information they capture. Sum communication in a team was measured as a total of all the interactions amongst team members. *Sharing task related information* refers to instances in which team members shared *agency specific information* and *actively sought to improve shared situational awareness*. Thus, whilst both the mediator and behaviour are related to "communication", the mediator focused on overall communication frequencies and the outcome focused on to what extent individuals shared task related information that only they had access to.

We also found that familiarity had an indirect impact on *conflicting priorities* through the increased frequency of communication within the team. This is interesting as other research has theorised the opposite: that failing to communicate can lead to conflict within MTS as component teams withhold information from one another (Hinsz & Betts, 2012). In contrast, the findings from the current Study suggest that familiar teams communicated more frequently, and this led to increased instances of *conflicting priorities*. The fact that *conflicting priorities* was positively correlated with positive behaviours suggests that it may have served a beneficial function in the team, such as enabling familiar teams to be more critical in task completion (*see* Shah & Jehn, 1993). Future research might consider reconceptualising the role of conflict in MTS as dependent on the familiarity/psychological safety present within the teams. Research has tended to present a negative view of conflict in MTS (*see* Cuijpers et al., 2016; Hinsz & Betts, 2012) and this may not be true of all teams.

The findings that communication, in part, mediates the effects of familiarity on team behaviours has important practical implications for MTS. They suggest that while increasing inter-agency training to improve familiarity amongst component teams ought to be a priority, targeted training to emphasise inter-team communication may have a similar impact. This is supported by research in conventional teams which suggests unfamiliar teams perform as well as familiar teams when information is shared effectively (Gruenfeld et al., 1996). In MTS, component teams will each have access to unique streams of information and if this is not shared effectively then it will not be possible to achieve shared situational awareness as each team will only have access to some of the relevant information. Thus, effective and open communication sharing across the system may alleviate some of the challenges MTS face when working with unfamiliar colleagues (*see* Fodor & Flestea, 2016; Waring et al., 2019). Future research might explore the effects of a communication-based training intervention on a unfamiliar MTS to ascertain if subsequent team behaviours are comparable to those in a familiar MTS.

Unlike for *sharing task related information* and *conflicting priorities,* there was little evidence that communication mediated the relationship between familiarity and *joint decision-making.* Findings were inconclusive as there was insufficient power to conduct mediation analyses, however we suspect there are other mediating factors, aside from the communication that may explain the relationship between familiarity and *joint decision-making.* Future research might consider other factors, such as the extent to which familiarity improves interpersonal processes in MTS (e.g., motivation and confidence building, *see* Killumets et al., 2015; trust and openness, *see* Alge, Wiethoff & Klein, 2003) and how this facilitates *joint decision-making.*

Limitations and future directions

Despite important theoretical and practical implications, this Study is not without its limitations. The first relates to the level of measurement within the Study. All team performance behaviours were measured at the MTS level; specifically, the extent to which team members engaged in the three positive and three negative coordinating behaviours (Brown & Power, 2018). To monitor the performance of a MTS as a whole, some have argued it is necessary to measure both MTS and component team efficacy (Asencio & DeChurch, 2017). For example, inter-agency communication at the MTS level might improve shared situation awareness across the network but lead to failures in intra-agency action as team members spend too much time focussing on inter-agency behaviours. Thus, behaviours may have a positive affect at the system level and the opposite effect at the component team level (DeChurch & Zaccaro, 2013). As our Study involved a simulated MTS in which there was only a single member of each component team represented, we were not able to measure intra- and inter- team performance. Whilst our Study has helped us to understand teamworking at the MTS level, future research may examine the variables considered here when multiple team members are present from component teams. This will allow us to establish if the system level effects generalise to a component team level.

The second limitation relates to the use of a student sample in re-creating a MTS. Simulation studies with students are common in the teamwork literature (e.g., Mathieu, Kukenberger, D'Innocenzo & Reilly, 2015), however, in this instance, the students were required to adopt the roles given to them to ensure that the simulation emulated a MTS. As we discussed earlier, this may have negatively impacted our ability to detect differences in *decision* and *role uncertainties* across familiar and unfamiliar teams, as students are not equipped with the necessary knowledge or experience of working in high-stakes settings as part of an MTS. With this aside, the use of students may have minimally affected the other study variables, as findings are consistent with prior research and familiar teams were recruited from sports teams in which we would expect to see established team processes and norms that would differentiate them from unfamiliar teams (Gruenfeld et al., 1996; Jehn & Shah, 1993; McEwan & Beauchamp, 2014). Moreover, as every effort was made to inform students prior to the simulation about the context of a major incident, the high level of experimental control afforded in this setting outweighs the difficulties of using a student population. Given the sample size (N = 22 teams) and the specificity of the context further research is needed to establish if findings apply to the MTS and in a practitioner-based sample.

Conclusion

Findings suggest familiarity improves MTS performance by increasing *joint decision-making* and the extent to which team members *shared task related information*. Counter to expectations, familiarity also increased instances of *conflicting priorities* in teams. This may indicate that conflict can be a positive team behaviour in instances where team members know one another and are able to employ cooperative conflict management strategies. Moreover, findings suggested the effect of familiarity is in part mediated by the frequency of communication within the teams. This provides some evidence that the benefits of familiarity can be partially achieved by encouraging more open and transparent communication across the MTS network. However, as this was not true for each of the relationships tested, findings also indicate that there are other processes (e.g., interpersonal processes, leadership) not considered in this Study that are likely important. Taken together, findings emphasise the need to increase inter-agency training for MTS responding in crises to overcome some of the difficulties of working within larger, complex systems.

Chapter V

Monitoring Team Cohesion over time in Expedition Teams; the role of Daily Events and Team Composition.

Chapter IV examined the role of familiarity in shaping multi-team system (MTS) behaviours. Findings showed that familiar teams made more joint decisions and shared more task related information than unfamiliar teams. These behaviours came about because familiar teams communicated with a much higher frequency than unfamiliar teams. This greater engagement in communication, however, also led to an increase in *conflicting priorities* (originally conceptualised as a negative team behaviour). This finding was attributed to psychological safety being present in familiar teams, enabling team members to question one another and voice their opinions without detrimental effects on performance (Edmonson, 1999; Roberto, 2002). This has potential implications for framing the effect of conflict in MTS as context-dependent, and while it may disruptive for some teams, it may benefit familiar teams. Taken together the results suggest there are benefits to increasing familiarity amongst team members and that organisations might achieve this through inter-agency training.

The central focus of this thesis is to understand how teams operate in extreme environments. In achieving this, this thesis makes an important distinction between types of extreme teams (ETs). Specifically, in *Chapter I*, I discussed how ETs can be differentiated as multi-team systems (MTS) who form quickly in response to emergencies and teams in isolated and confined environments (ICE), who operate for longer periods in inhospitable climates. Whilst the literature on ETs has tended to focus on drawing out the challenges and characteristics of these teams in comparison to conventional teams, it rarely addresses the challenges presented to different types of ETs. Both empirical Chapters thus far (*Chapter III* and *Chapter IV)* have focused on generating a better understanding of teamwork in MTS responding in crises. This Chapter focuses on teams in ICE, serving as a comparison point to the prior research on MTS, highlighting differences in the importance of variables and the appropriateness of methodological approaches.

As discussed in the previous Chapters, teamwork is difficult in MTS as multi-agency teams are formed quickly, there can be a lack of familiarity and agreement of objectives amongst teams and teams are required to maintain effective communication channels across larger, often disparate networks (Fodor & Flestea, 2016; Waring, Alison, Shortland & Humann, 2019). These challenges led to a focus on communication, coordination and familiarity in the preceding Chapters. In contrast, the primary challenges to teams in ICE relate to the harsh physical environment and how to maintain effective teamwork in continued close contact with a small number of individuals (Barrett & Martin, 2014; Roma & Bedwell, 2017). Whilst I am not suggesting communication, coordination and familiarity will not be important for teams in ICE, there has been an emphasis on the importance of cohesion in teams who exist for longer periods (Vessey & Landon, 2017). This is because the environment in which teams in ICE operate is expected to amplify the cohesion-performance relationship (Stuster, 2011). Further, as conflict in ICE is more likely to happen as a result of individual differences, the personality traits of team members are an important consideration (Bell & Outland, 2017; Stuster, 2011). In addressing the first question of this thesis, "*What factors support effective teamwork in different types of extreme environments?*" the primary aim of this Chapter was to identify how cohesion emerges and is sustained in ICE and to explore if this is influenced by the personality composition of the teams.

In answering the second question of this thesis, "*What methods and analytical approaches are suitable for studying teams in different types of extreme environments?*", the secondary aim of this Chapter was to test the utility of a diary-based methodology in studying teamwork in ICE. Thus far, each chapter has focused on the use of immersive simulations as a way of studying teams in organisational settings (*see Chapter II)*. Whereas *Chapter III* generated a greater understanding of how emergency responders manage teamwork in MTS by collecting data from a training exercise with practitioners, *Chapter IV* provided an opportunity to test specific aspects of MTS theory within the laboratory. Simulations have been used to study teams in ICE, however, this has tended to be for those involved in long-distance spaceflight (Sandal, Bye & Van de Vijver, 2011). For other teams in ICE, such as those on expedition, anti-poaching units and defence and security personnel, simulations are unlikely to be of use as they cannot re-create the ever-changing landscape in which these teams operate. Diary methods offer a solution of this. Diary methods have been used widely to research teams in ICE (Atlis, Leon, Sandal, & Infante 2004; Smith, Sandal & Barrett, 2018), however, they tend to focus on *individual* performance and adaptability. In extending prior research, this Chapter focuses on how diary methods can be utilised to monitor team cohesion.

Pre- and post- expedition questionnaires were distributed to *n* = 68 participants. Diaries were completed by *n* = 42 participants, forming five teams taking part in expeditions lasting twenty days. Analysis used pre- and post-expedition data to identify how team composition influenced changes in cohesion over time. Liner mixed model analysis of the diary data was then used to predict which factors influenced fluctuations in cohesion over time.

Abstract

Cohesion is an important part of effective team performance. Previous research has focused on cross-sectional self-report measures in business settings. However, in extreme environments where contextual factors (e.g., weather conditions) can vary considerably from day to day, micro-variations in cohesion could influence daily performance. In small teams under pressure, such variations may be moderated by personality traits. The current study presents a diary methodology to explore variation in cohesion in five expedition teams – tracking temporal changes in cohesion and daily events over twenty days. Pre-expedition personality measures were used to explore the impact of team composition on variations in cohesion. Findings demonstrated that events significantly predicted fluctuations in cohesion across teams. Having more extraverted team members had a negative impact on cohesion. These results offer valuable insight to how this method can track changes in cohesion over time and subsequently enhance understanding of how to mitigate cohesion breakdowns.

Keywords: *Teamwork, Cohesion, Team Composition, Performance*

Monitoring Team Cohesion over time in Expedition Teams; the role of Daily Events and Team Composition

Expeditions teams operate in extreme, high stakes, pressured environments in which effective teamwork is vital to success (Militello, Sushereba, Branlat, Bean & Finmore, 2015; Driskell, Salas, Driskell, 2018). Research on how teams work effectively in extremes is an important, emerging area of research (Salas, Tannenbaum, Kozlowski, Miller, Mathieu & Vessey, 2015). It is not yet known if facets of teamwork will operate in extremes in the same way they do in other environments (Vessey & Landon, 2017). Despite this, there are only a small number of studies that have empirically tested teamwork in extremes and even fewer with a longitudinal design, accounting for the dynamic nature of team interactions. The primary purpose of this research was to trial a method for monitoring team cohesion over time in expedition teams and assess if changes in cohesion could be explained by daily experiences. In contrast to previous research, which has tended to collect data from a single expedition team (e.g., Allison, Duda & Beuter, 1991), we collected data from five teams, travelling to three different locations. This allowed a comparison of how team composition affected the development of cohesion in each team. The majority of participants were high school students, taking part in expeditions lasting 20 days, travelling to three different countries; Mongolia, Kyrgyzstan and Greenland.

Expeditions are defined as journeys, taken with purpose for reasons of adventure, exploration and scientific discovery (Johnson, Anderson, Dallimore, Winser, & Warrell, 2008). Generally, when on expedition, teams tend to be socially isolated and physically confined, in environments characterised by dangerous and difficult conditions (Palinkas & Seudfeld, 2008). This isolation, in combination with the arduous physical demands associated with expeditions, makes establishing and maintaining effective teamwork difficult, increasing the likelihood of social conflicts and exaggerating individual differences (Palinkas & Suedfeld, 2008; Stuster, 2011). Monotonous daily tasks (e.g., setting up and taking down camp) are also typical of an expedition environment (Leon, Kanfer, Hoffman & Dupre, 1994), increasing the likelihood of conflict as individuals contend with feelings of boredom. Despite the challenges to maintaining teamwork on expedition, it is a vital component to success. Failing to work effectively will lead to suboptimal decision-making with consequences for the safety and performance of the team (Driskell et al., 2018). Teams must be able to coordinate, communicate and cooperate effectively, relying on each other to cope with the challenges of the environment (Bishop, Morphew & Kring, 2000). On reflection of a crossing of the Arctic one member noted: "If you don't have the team you have nothing. Have team members who have social intelligence… anyone can learn tasks" (Leon, Sandal, Fink & Ciofani, 2011, p.14).

To examine teamwork in expedition teams, we adopted the Input Mediator Outcome Input (IMOI) model of teamwork (Ilgen, Hollenbeck, Johnson & Jundt, 2005). This model accounts for the dynamic nature of team interactions, acknowledging that teams exist within the wider environment, changing over time (Kozlowski & Ilgen, 2006). According to this model, the relationships between Inputs (e.g, team size, team composition) and Outcomes (e.g., performance measures) occur in a cyclical process, influenced by Mediators (e.g., processes and emergent states). Team processes are defined as team interactions that are directed towards task accomplishment (Mathieu, Maynard, Rapp & Gilson, 2008) and emergent states are defined as the dynamic properties of the team representing attitudes values and cognition (Marks, Mathieu & Zaccaro, 2001). To study teamwork in expedition teams, we focused on team composition (an input) and team cohesion (an emergent state). Both have previously been identified as important aspects of performance in extremes (Bell & Outland, 2017; Vessey & Landon, 2017).

Cohesion is defined as the shared tendency for the team to remain united in achieving a common goal (Casey-Campbell & Martens, 2009). McClurg, Chen, Petruzelli and Thayer (2017) formulate it more simply as the commitment of a team to work on a task constructively, whilst maintaining social relationships. Cohesion tends to be viewed according to its task and social components (Boyd, Kim, Ensari & Yin, 2014; Carless & De Paola, 2000). Task cohesion is defined as the shared commitment to the task, and social cohesion is defined as the interpersonal bonds that exist between team members (Mikalachki, 1969). In conventional teams, cohesion has been consistently associated with high performance (Beal, Cohen, Burke & McLendon, 2003) and has been found to facilitate team decision-making under pressure (Zaccaro, Gualtieri & Minnionis, 1995). Despite this, little research has explicitly measured cohesion in expedition teams, and most existing studies have only used a single item to measure it. Data from a team of 12 members, completing a 61-day trek through parts of Alaska, identified a positive association between cohesion and communication, perception of fairness in task assignments and in the perceived quality of decision making by the team leader (Leon et al., 1994). A further study of an all-female climbing group found that cohesion increased during the early part of the expedition, peaking on the day when the group engaged in their most difficult task, before tailing off towards the end (Allison et al., 1991). Our research built on the methodology used by Allison et al. (1991), however it utilised a validated measure of cohesion, collected at daily intervals, rather than at six pre-defined intervals.

Team composition is defined as the attributes of team members, including skills, experiences and personality characteristics (Guzzo & Dickinson, 1996). Composition is an input factor, relating to the extent that attributes of team members effects emergent states, processes and outcomes of teams (Mathieu et al., 2008). Specifically, for expedition teams, composition has been identified as an important way of screening those most suited to survival in harsh environments and to support the achievement of team goals (Palinkas & Suedfeld 2008). Previous research has identified traits of openness to new experiences, agreeableness and conscientiousness in expeditioners (Suedfeld & Steel, 2000; Steel, Suedfeld, Peri & Palinkas, 1997). Palinkas, Gunderson, Holland, Miller and Johnson (2000) sought to explore predictors of performance in extremes by identifying the traits of 657 men overwintering in Antarctica. The results of their study posited that low levels of extraversion would be beneficial in an extreme environment. The authors suggest that this may be due to the restrictive social environment of isolated contexts being more suited to less extraverted individuals. Despite identifying common and beneficial traits in expedition teams, studies have not yet compared the composition of several teams and how this composition might interact with other aspects of teamwork (i.e., cohesion). Of the limited research that has been conducted, in conventional teams an association has been reported between emotional stability and cohesion (Barrick, Stewart, Neubert & Mount, 1998) and agreeableness and cohesion (Bradley, Baur, Banford & Postlethwaite, 2013).

The primary purpose of this research was to pilot a diary methodology for monitoring team cohesion over time in expedition teams and assess if changes in cohesion could be explained by daily experiences. We expected that in small teams, changes in cohesion might be moderated by the personality traits of team members.

Thus, the second aim of our research was to compare the composition of each of the expedition teams and explore if the team composition affected change in team cohesion.

Method

Participants

A total of 71 participants (43 of whom were female) were recruited from a school in the south of England. Participants formed five teams of varying size, travelling to three locations; Greenland, Mongolia and Kyrgyzstan. Each expedition lasted 20 days. The teams included staff members (*n*= 9) and students (*n* = 62). The average age of the student participants was 15.22 years (*S.D* = .35). 68 participants completed the pre-expedition questionnaire, 42 participants completed the diary and 50 completed the post-expedition questionnaire. Ethical approval for this study was granted by the Faculty of Science and Technology Research Ethics Committee, Lancaster University (FST16125).

Procedure

Following ethical approval, contact was initiated with the school several months before the expedition teams were due to depart. Once a formal agreement was made between the researchers and the school, a letter of consent was sent out to the parents of the students taking part in the expedition. Once consent was obtained, participants were briefed on how to complete the daily diary. Each participant was asked to complete a pre-expedition questionnaire and given a daily diary to complete on each day of the expedition. On return, participants were asked to complete a post-expedition questionnaire.

Materials

The pre-expedition questionnaire included questions about demographics (age, gender), a personality measure and a team cohesion measure. Personality was measured with the Ten Item Personality Measure (TIPI) (Gosling, Rentfrow & Swann, 2003), a short measure designed for instances when time is limited. The TIPI measures the Big-Five Personality Traits: Extraversion ($\alpha = .50$), Openness to new experiences ($\alpha = .45$), Conscientiousness ($\alpha = .45$), Agreeableness ($\alpha = .35$) and Emotional stability ($\alpha = .77$)[4]. Cohesion was measured using a 6-item scale (Mathieu, 1991), with three items measuring social cohesion ($\alpha = .84$) and three items measuring task cohesion ($\alpha = .85$). This scale has been used to assess the impact of shared leadership and team members competence on team cohesion and performance over time (Mathieu, Kukenberger, D'Innocenzo & Reilly, 2015). The daily diary was designed to mirror those used in previous research (Smith, Barrett & Sandal, 2018). It included 27 diary items (*see* table 2), with 18 negative items (e.g., *"delay due to weather conditions")* and 9 positive items (e.g., *"enjoyment of the environment")*. Participants were instructed to place a tick next to each event they experienced that day. The diary also included the same cohesion measure (Mathieu, 1991) as the pre-expedition questionnaire and a single item measuring perceptions of team performance. In the post-expedition questionnaire, participants were asked to complete the cohesion measure again.

Data Analysis

No differences were identified in the scores of staff members and student participants in any of the measures used in the study (avg. $t = -.14$, $p = >.05$).

[4] The low internal consistency estimates are consistent with other reported uses of this scale and are attributed to the fact that the scale measures broad domains with only 10 items. It is a measure designed for instances where personality is not the sole topic of interest (Gosling et al., 2003) (https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/a-note-on-alpha-reliability-and-factor-structure-in-the-tipi/)

Accordingly, scores from each group were collated for the analyses. Differences in personality traits across the teams were explored using a MANOVA. A mixed ANOVA and Bayesian statistics were used to examine if cohesion increased following the expedition and if this varied according to group. A linear regression was conducted to explore the extent to which change in cohesion could be explained by the personality composition of the team. To analyse the diaries, R (R Core Team 2012) and lme4 (Bates, Maechler & Bolker, 2012) were used to perform a linear mixed model analysis to explore which variables significantly predicted changes in cohesion. A major challenge to researching teams in extremes is that sample sizes tend to be small (Bell et al., 2018). However linear mixed model analysis is able to produce reliable results with relatively small samples (Bell, Morgan, Schoenberger, Kromrey & Ferrron, 2014). As fixed effects, the 27 diary items were included along with the participant's team (e.g., Greenland) and the day of the expedition the diary was completed (day 1-20). As random effects, an intercept was added for participants. Intercepts did not vary according to the team participants were in, thus team was included only as a fixed effect. A visual inspection of residual plots confirmed there was no major deviations from homoscedasticity or normality. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

Results

Questionnaire: Personality composition

The mean personality profiles for each team were calculated by aggregating the individual personality data collected in the pre-expedition questionnaire. This is common practice in team research (Mathieu, Gallagher, Domingo & Klock, 2019). A MANOVA was conducted to compare the personality profiles in each team, no significant differences were identified. The most frequent trait found in the sample was openness to new experiences, followed by conscientiousness and agreeableness. The lowest scoring traits were emotional stability and extraversion (*see* Table 1).

Table 1. Personality profiles of each team

Team	Extraversion	Agreeablenes	Conscientiousness	Emotional stability	Openness to new experience
Mongolia 1	4.87	5.10	4.88	5.72	5.31
Mongolia 2	4.96	4.90	5.54	4.75	5.18
Kyrgyzstan 1	4.08	4.80	5.41	5.08	5.62
Kyrgyzstan 2	4.73	5.11	5.38	4.92	5.73
Greenland	4.08	5.20	5.31	4.88	5.70
Total	**4.74**	**5.022**	**5.29**	**4.86**	**5.50**

Questionnaire: Cohesion scores before and after the expedition

There were no significant differences between social and task cohesion scores in the pre- and post-expedition questionnaires, $t(95) = -1.18$, $p = .24$. Due to this, analysis was only conducted to compare the total cohesion scores before and after the expedition. A two-way mixed ANOVA was conducted to explore differences in team cohesion across the five teams and before and after the expedition. Results indicated a significant main effect of time, $F\ (1{,}86) = 21.25$, $\eta p^2 = .20$, $p = <.001$, $B_{H(0,\ 2559.31)} = 0.55$[5]. with overall cohesion significantly higher after the expedition (M=6.30, $S.D$ = .67) than before (M = 5.70, $S.D$ = 6.70). There was no significant difference in cohesion across the groups. There was however a statistically significant interaction between the increase in cohesion and the expedition team $F(4{,}86) = 3.75$, $\eta p^2 = .15$, $p = .007$. A simple main effects analysis found that cohesion increased significantly over time for Mongolia team 2 (p=.019, $B_{H(0,\ 5.37)}$, = 0.55) and the Greenland team (p = <.001, $B_{H(0,}$

[5] The alternative hypothesis was set by the mean difference reported by Chang and Bordia (2001) when exploring changes in cohesion over time.

$_{36.75)}$ = 0.55), but not for the other three remaining teams (*see* Figure 1. Error bars included for Greenland and Mongolia 2).

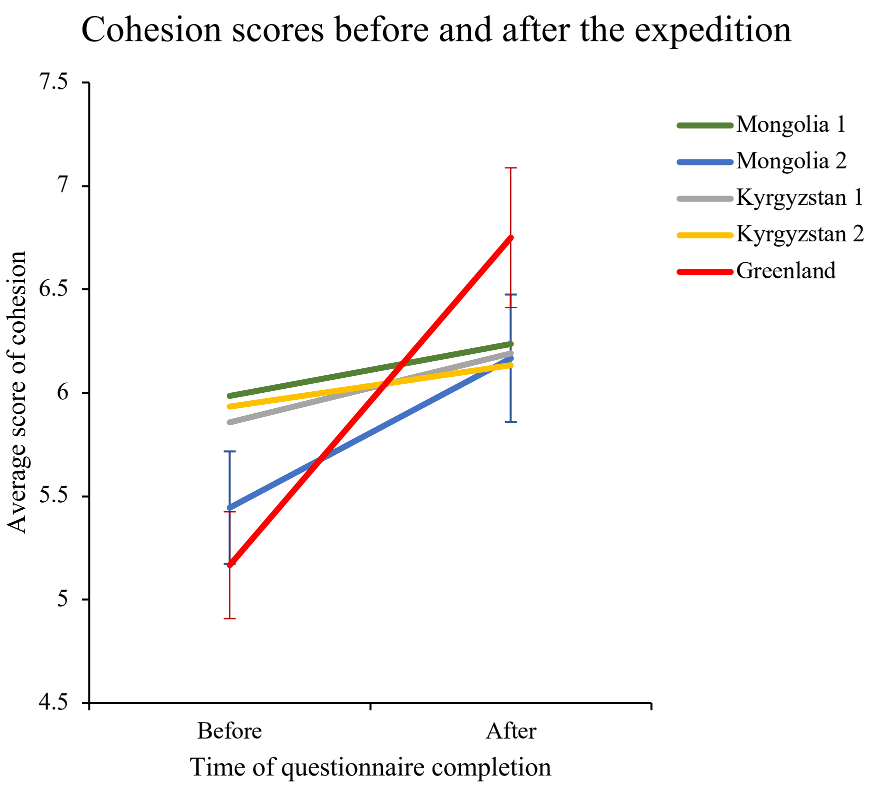


Figure 1. Team Cohesion over time

Questionnaire: Personality composition and change in cohesion

A linear regression was conducted to explore if the change in an individuals' perception of cohesion could be explained by the aggregated personality scores for each team. The overall model was significant (F(4,43) = 4.316, p = .005), $R^2_{adjusted}$ of .22, indicating 22% of the variance in cohesion change could be explained by the personality composition of the teams. However, of each personality trait, only extraversion significantly predicted a change in cohesion (β = -.53, p = .002).

Diary

The most frequently reported diary item was *"Enjoyment of Environment" (n* = 608) and the least frequent item was *"Dispute with the leader" (n* = 54). The findings were consistent across the teams. Despite there being twice as many negative items in the diary (*n* = 18) as positive items (*n* = 9), positive items were reported much more frequently (59.4% of all responses). Whilst there were differences between teams in perceptions of cohesion in the pre- and post-expedition questionnaires, there was no significant effect of team in predicting scores of cohesion *during* the expedition, $\chi 2$ (1) = 1.74, p = .19. The diary data was therefore collated across teams to identify if day of completion or content of the diary could predict changes in cohesion (*see* Table 2). Perceptions of cohesion significantly correlated with perceptions of performance *r* (808), = .81, p = <.001 and had an unacceptable level of collinearity to be included in the mixed model analysis. The linear mixed model analysis demonstrated that the day that the diary was completed had a significant effect on cohesion, $\chi 2$ (1) = 4.59, p = .03, indicating that as the number of days increased, cohesion increased by .06 ± 0.06. The results of the linear mixed model analysis with regards to the utility of the diary items in predicting changes in cohesion are presented in Table 2. Interestingly "*Satisfaction in making good progress today"* and *"Satisfied that I am able to cope with the challenges of the expedition"* were significantly associated with increases in cohesion. In contrast *"Problem with digestion"* was associated with decreases in cohesion. Perhaps not surprisingly *"Feeling of camaraderie/closeness with team mates", My team mates approached the expedition today with a good attitude",* "*Feeling I could rely on my team mates to work effectively*" were all associated with increases in cohesion, whereas "*Concern about how effectively my team and I are working together", "Feeling down/low because my team mate is/are feeling that way"* and

"Tension or argument with team mate(s)" were all significantly associated with decreased cohesion.

Table 2. Utility of daily events in predicting changes in cohesion over time

Daily event	Frequency of item	Significance	Change in cohesion
Problems with gear and equipment e.g., clothing, tools, navigation equipment etc.	164	ns	
A delay in progress due to weather conditions	129	ns	
Worried about encountering bad weather	216	ns	
Enjoyment of the environment	608	ns	
Satisfaction that equipment is working properly	423	ns	
Satisfaction in making good progress today	427	.032	+.15
Satisfaction that I am able to cope with challenges	440	.039	+.11
Concerns about the effectiveness or safety of the decisions I made today	86	ns	
Concern about the well-being of my team mates	251	ns	
Tension or argument with my team mate(s)	191	.027	-.13
Discussed a problem with a team mate and felt listened to	269	ns	
Feeling of camaraderie/closeness with team mates	529	.013	+.14
Feeling down/low stressed out because my team mates is/are feeling that way	91	.02	-.19
Feeling I could rely on my team mates to work effectively	423	<.001	+.25
Concern about how effectively my team mates and I are working together	89	<.001	-.43
My team mates approached the expedition today with a good attitude	574	<.001	+.25
Satisfaction with the leadership	447	ns	
A problem/dispute with the leader	54	ns	
Problem with digestion	65	.016	-.20
Headache	87	ns	
Lack of sleep	318	ns	
Muscle or joint pain	206	ns	
Personal hygiene (wanting to be cleaner)	278	ns	
Lack of privacy/personal time	163	ns	
Fear of being injured	119	ns	
Loneliness, homesickness	175	ns	
Worried about family/friends	110	ns	

Discussion

The overall purpose of the present study was to trial the use of a diary methodology to examine teamwork in five teams, undertaking expeditions in three locations. We wanted to understand if the composition of the team and the daily experiences of individuals in each team could predict changes in cohesion. Despite no significant differences in personality composition across the teams, the highest scoring traits (openness to new experiences, conscientiousness and agreeableness) were consistent with traits that have previously been identified in expeditioners (Smith, Kinnafick, Cooley & Sandal, 2017; Suedfeld & Steel, 2000), adding to the existing evidence that these personality traits are most suited to expedition environments (Palinkas & Suedfeld, 2008).

In comparing the pre and post-expedition questionnaires an overall increase in cohesion was found. Post-hoc tests indicated that cohesion only significantly increased for two of the teams. This finding supports the view that cohesion emerges and changes over time depending on the dynamics in the environment and the team (Marks et al., 2001). By demonstrating differences in the emergence of cohesion across several teams, one theoretical contribution of this research is to support the notion of context being vital in shaping team-based constructs (Ilgen, 1999). If cohesion were to increase organically in teams regardless of context, we would expect to see a similar change across each of the teams, however in our findings an increase was only found in two of the teams. To further explore this, we assigned the aggregated team personality profile to each individual, to see if personality composition could explain the changes in cohesion scores before and after the expedition. The results indicated that personality accounted for a fifth of the variance in scores of cohesion. Further analyses demonstrated that extraversion was the only trait to have a significant effect, leading to a reduction in cohesion. Findings therefore suggest having extraverted team members had a negative impact on the development of cohesion. This provides a possible explanation for why previous research identified low levels of extraversion as beneficial in an extreme environment (Palinkas et al., 2000). Overall the results from the questionnaire support previous suggestions that certain traits may be more beneficial for performance in extreme and challenging environments. By identifying extraversion as an important variable in the emergence of cohesion, the findings support other research suggesting that compositional factors, such as personality, can affect team processes and emergent states (Bell & Outland, 2017).

Whilst on expedition, participants completed a diary, monitoring their experiences of daily events and perceptions of team cohesion and performance. Participants consistently reported more positive diary items than negative, and the most frequently reported item across each of the teams was "*Enjoyment of the environment*". This finding supports existing evidence that extreme environment activities can be promotive of health (Suedfeld & Steel, 2000) and a positive experience for those taking part (Smith et al., 2018). Consistent with findings in the literature we found a significant positive relationship between cohesion and performance. This is the first study of its kind to test the relationship between cohesion and performance in an expedition setting and it is promising to see evidence for the importance of cohesion in this context.

In addition, consistent with the results of the questionnaires, we found a significant relationship between the day the diary was completed and perceptions of cohesion. As the number of days increased, so did scores of cohesion. Whilst there are very few studies that have studied cohesion over time in expedition teams, this finding is inconsistent with one example of an all-female climbing group, in which cohesion was found to decline towards the end of the expedition (Allison et al., 1991). The authors suggested that the decline in cohesion may have been because the team had already achieved their main goal, leading to attention shifting from the team towards matters at home. As the participants did not know each other before the expedition, this then led to personal investment in the team waning. In our study, the participants had trained together for two years prior, which might explain why increases in cohesion sustained throughout.

The diary method allowed us to test the validity of monitoring daily events to predict fluctuations in cohesion. Similar methods were used by Smith et al. (2018) to predict fluctuations in positive and negative affect by monitoring daily events and coping strategies. Some associations between the team-focused events and cohesion were expected. For example, "*camaraderie with teammates*" led to an increase in perceived cohesion and "*feeling concerned about the effectiveness of the team*" led to a significant decrease in perceived cohesion. This validates the cohesion measure as reflective of changing perceptions of teamwork across the expedition. Other associations that were made between the daily events and perceptions of cohesion were less obvious and may be of particular value in indicating how to mitigate breakdowns in cohesion. Reporting "*feeling satisfied to cope with the challenges of expedition*" and "*feeling satisfied with the progress of the expedition*" both led to increases in cohesion. These results suggest that experiencing a sense of achievement (satisfaction of coping with challenges) and achieving shared goals (such as progressing in the expedition) are important aspects in the maintenance of team cohesion. Previous work has theorised how this process occurs by suggesting that superordinate team goals encourage social identity, which is an important component in the development of cohesion (Salas & Cannon-Bowers, 2001).

Only one diary item that was not explicitly related to teamwork significantly contributed to a reduction in cohesion and this was reporting a problem with digestion. Problems with digestion and diet have previously been identified as major stressors, leading to increased tension amongst crew mates during a 105-day space simulation (Sandal, Bye, Van de Vijver, 2011). Overall the results from the diary data are promising, demonstrating the benefit of the diary methodology to track cohesion over time and assess the utility of daily events to predict changes in cohesion. This was an exploratory study; further research should explore if the relationships between variables identified in this study are consistent for teams operating in other contexts. This could facilitate the development of a mobile monitoring system, allowing team leaders to monitor fluctuations in cohesion in real time and therefore mitigate breakdowns in effective teamworking.

Conclusion

Despite being an exploratory study, the findings are promising and identify the need to conduct further research in this context. We successfully piloted a daily diary method to understand how cohesion is established and maintained in difficult conditions, allowing us to monitor changes in cohesion across the entirety of the expedition, assessing to what extent these changes could be explained by daily events and team composition. These findings offer valuable information to expeditioners and other teams operating in analogous settings (e.g., aid workers, special forces personnel) on factors that might influence cohesion, as well as contributing to the theoretical understanding of how cohesion emerges in different contexts.

Chapter VI

Discussion

Discussion

In this Chapter I summarise the findings across the three empirical chapters, consider the theoretical and practical implications from my thesis, highlight the limitations to my work and suggest avenues for future research in the study of ETs.

This thesis examines teamwork in extreme environments. Increasing attention has been paid to these environments and the teams that operate within them (Burke, Shuffler & Wiese, 2018; Vessey & Landon, 2017). With this has come a consideration of the methods and analytical techniques suited to researching these complex, hard to reach extreme teams (ETs; Kozlowski, 2015). This interest has led to Special Issues of journals (*see* Maynard, Kennedy & Resick, 2018), review papers (Driskell, Salas & Driskell, 2018; Golden, Chang & Kozlowski, 2018) and guidelines for conducting research (Bell, Fisher, Brown & Mann, 2018). However, there remains a lack of empirical research that has sought to understand what supports effective teamwork in extremes and even less that has differentiated between *types* of extreme environment. In light of this, this thesis presents a series of empirical studies on two distinct types of ET: emergency response teams and expedition teams. Findings emphasise the importance of considering context-specific challenges to guide empirical research in ETs and add to the limited evidence base seeking to understand how to promote effective, safe working in these unique, often hard to reach teams.

Revisiting the research questions

RQ1: What factors support effective teamwork in different types of extreme environments (MTS and ICE)?

In answering RQ1, I first reviewed the literature on two types of ET: multi-team systems and teams in isolated and confined environments (*Chapter I*). A central contribution of this review was showing how, despite most researchers identifying ETs as being distinct to conventional teams (*see* Driskell et al., 2018; Maynard et al., 2018), they fail to consider intra-extreme differences. Despite all ETs working in atypical environments where failure can have severe consequences, these teams can be differentiated on a number of features (Bell et al., 2018). For example, some teams combine with other agencies forming multi-team systems (MTS) that respond quickly in crises, whereas other ETs exist for longer periods (e.g., weeks or months) in isolated and confined environments (ICE) (*see* Figure 1 for defining features of each type of ET).

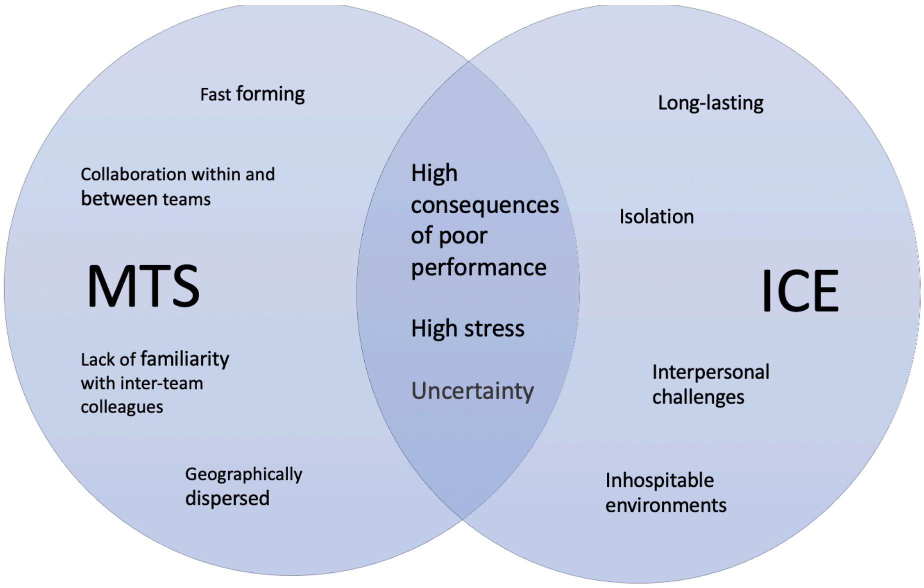


Figure 1. Typology of intra-extreme differences.

An important implication of the review in *Chapter 1* is that delineating between MTS and ICE may better identify the factors of teamwork that are relevant to performance in each context. In reviewing the current literature on MTS and ICE, it became clear that there were differences in the importance of teamworking variables. For example, cohesion is cited as being essential to effective teamwork in ICE as it reflects the ability of team members to sustain social relationship whilst remaining focused on the task (McClurg, Chen, Petruzelli & Thayer, 2017; Stuster, 2011). However, for MTS, intra-team cohesion may derail system-level working by intensifying divisions and fuelling a competitive atmosphere across the system (DiRosa, 2013). For MTS, it is more important that component teams are able to come together and quickly coordinate actions, as there will often not be time for cohesion to emerge and develop in the short period in which they are working together (e.g., in crisis response). As such, no single variable was examined in both contexts. I did, however, consider personality effects across contexts (see later).

In studying MTS, communication, coordination and familiarity were identified as important for further empirical study. These variables speak to the key challenges to teamwork in MTS, which relate to: (i) the need to quickly bring together multi-agency teams; (ii) establishing communication channels across geographically dispersed, disparate teams; (iii) balancing the goals and priorities of multiple teams in parallel, and (iv) an absence of familiarity between inter-agency colleagues (Crichton, Flin & Rattray, 2000; Fodor & Flestea, 2016; Luciano, DeChurch & Mathieu, 2018; Waring, Alison, Shortland & Humann, 2019). For teams in ICE, cohesion and individual differences were identified in the literature as being important factors to consider in empirical work. These variables map onto the key challenges to teamworking in ICE: (i) the need for teams to remain cohesive so that they are able to live *and* work alongside each other, and (ii) prolonged contact with a small number of individuals exacerbating individual differences (Stuster, 2011; Vessey & Landon, 2017).

An empirical examination of the factors identified in the literature review of *Chapter I* supported their proposed importance for effective teamwork in ETs. The thesis offered empirical support for using recent theoretical advancements within the

MTS literature to interpret findings. With relation to MTS, it also examined how communication and coordination change over time (*Chapter III)* and if they are influenced by familiarity (*Chapter IV)* amongst team members. One of the most significant contributions of this thesis was the identification of specific classes of coordinating behaviours (behaviours that enabled team members to synchronise their efforts to achieve goal related outcomes). This is the first set of such behaviours to be identified and we used these as our "outcome" indicators in both studies on MTS (Mathieu, Maynard, Rapp & Gilson, 2008; Marks, Mathieu & Zaccaro, 2001). Researchers agree coordination is vital to support effective teamwork in MTS (e.g., Rico, Hinsz, Davison & Salas, 2018), however, little is known about the behaviours that underpin it (Wijnmaalen, Voordijk & Rietjens, 2018). This thesis addresses this absence by identifying behaviours that enabled MTS to synchronise their efforts to achieve goal related outcomes (positive coordinating behaviours) and those that disrupted their efforts (negative coordinating behaviours). Transcripts of team member interactions were thematically analysed, identifying three positive coordinating behaviours: *joint decision-making, sharing resources*, and *sharing task related information*, and three negative coordinating behaviours: *conflicting priorities, role uncertainties,* and *decision uncertainties*. These sets of behaviours were identified across two studies with different populations.

The findings of *Chapter III* showed that coordinating behaviours, and indeed communication, varies across time. These results were one of the first to empirically show the importance of considering temporal changes in teamwork in MTS. They suggest that understanding what factors support effective teamwork in ETs may vary dependent on time, and that behaviours are influenced by changing contextual demands. For example, *conflicting priorities* (i.e., the extent to which agencies attempted to re-orient conversations to suit their own intra-agency priorities) occurred more frequently in the initial phase of emergency response. In this phase, agencies had very different tasks (e.g., Police – neutralising the threat, Ambulance – reaching and prioritising causalities), whereas in latter phases agencies were all focused on business continuity and re-establishing community cohesion and there were no instances of *conflicting priorities*. This finding provides empirical support for Luciano et al.'s (2018) theoretical framework of MTS. Within this framework, teams are said to be differentiated by work *process dissonance*, that is, the extent to which team members are working separately on different tasks as opposed to interdependently. When process dissonance is high, team processes are proposed to be impeded, which the results in *Chapter III* of an increase in *conflicting priorities* supported.

Interestingly, whilst *conflicting priorities* disrupted teamwork in the simulation study with emergency responders (*Chapter III),* they increased with positive coordinating behaviours in the lab-based study (*Chapter IV*) and were more frequent when team members were familiar with one another. Familiarity appeared to alter the way in which *conflicting priorities* operated, suggesting it may have had a positive effect on performance by enabling team members to be more critical in task completion (*see* Shah & Jehn, 1993). This finding may be attributed to familiar teams sharing a sense of psychological safety; a confidence to speak up and to challenge ideas without fear that it will be used against them (Edmonson, 1999; Roberto, 2002). Thus, whilst *conflicting priorities* was identified as a negative behaviour that may disrupt coordination in MTS, the findings suggest that familiarity amongst team members can temper this. This is an important finding and in contrast to previous research which has tended to view conflict as disruptive to team performance in MTS (Hinsz & Betts, 2012). It would instead seem that this relationship is more complicated and that under certain conditions *conflicting priorities* may benefit team performance. This is reflected in the conventional team literature (DeChurch & Marks, 2001; O'Neill & McLarnon, 2018) and suggests more research is needed to identify under what conditions (in addition to teams being familiar) conflict can be of benefit to MTS performance.

Both studies focusing on MTS supported the finding that coordination and communication are inextricably linked (Rosen, Fiore, Salas, Letsky & Warner, 2008). The findings reported in this thesis show that a team that communicates well will likely employ effective coordinating behaviours in task completion. For example, in *Chapter III* we found improvements in team behaviours across time coincided with networks becoming less centralised. This suggests teamwork improved as a larger number of agencies became involved in discussions, supporting prior research that decentralised networks are of benefit to MTS during emergencies (Schraagen & Van de Ven, 2011; Foder & Flestea, 2016). In *Chapter IV*, centralisation was not related to any of the coordinating behaviours. Instead, frequency of communications was strongly related to positive team behaviours. This is an important finding as some researchers have argued against using frequency counts of behaviour as an indicator of performance (Stachowski, Kaplan & Waller, 2009). This is because some teams share less information because they have a greater implicit knowledge of one another's working and not because they are deficient in communicating (Stout, Cannon-Bowers, Salas & Millanovich, 1999). The findings in this thesis suggest that this relationship may be more nuanced. When teams are brought together quickly (i.e., in MTS), then frequency of communications may be a useful indicator of performance, particularly in smaller and co-located MTS as shown here. When teams work together on a regular basis (i.e., in command and control teams, *see* Waller, Gupta & Giambatista, 2004), then frequency of communications may be a poor measure of performance as team members are more likely to engage in implicit forms of communication.

The number of members in a MTS may explain why different measures of communication (e.g., centrality or frequency) matter more to performance. In *Chapter IV,* teams comprised 5 members who all spoke to each other at least once during the simulation. Accordingly, a weighted centrality measure was used (eigen vector centrality) to account for the relative distribution of communication across team members. This is different to the centralisation measure used in *Chapter III,* which measured how connected the network was, given that many team members (there being 28 members in total) did not interact. Taken together, findings suggest that in smaller, co-located MTS in which each team member is in contact with one another, the total amount of communication across the system will be more important than the extent to which everyone contributes an even amount (eigen vector centrality). As long as key information is shared (as evidenced by high levels of *sharing task related information*), it may not matter if certain agencies contribute more than others as each agency will remain informed of all information shared.

However, in larger and more diversified MTS (as in *Chapter III)*, in which it can be assumed that certain agencies will not interact, the centralisation of networks appears to be more important. In this instance a centralised network will result in many agencies not contributing at all, leading to the loss of their expertise and a failure to consider all of the information available. Thus, in larger teams, decentralisation is important to ensure that agencies are engaged in the response and information can be distributed throughout the network (Schraagen & Van de Ven, 2011). In answering RQ1, findings indicate that communication is vital to support effective teamwork in MTS. However, it is important to also account for the size and diversity of the MTS as this will likely influence the relative importance of network centralisation or frequency of communications across the system in driving team behaviours.

One of the reasons for continued challenges to teamwork in emergency response teams is a lack of familiarity amongst team members from different agencies (Waring et al., 2019). The findings in *Chapter IV* showed how familiarity may alleviate some of the challenges associated with working in larger multi-agency teams by demonstrating how familiar teams made more joint decisions and shared more task related information than unfamiliar teams. Findings were consistent with prior research on conventional teams (Espinosa, Slaughter, Kraut & Herbsleb, 2007; Huckman, Staats & Upton, 2012). Further analysis examined how familiarity improved teamwork by assessing the extent to which this occurred due to communication structures and shard goals. The results showed differences in behaviours were not a result of increased shared goals in familiar teams. However, partial support was found for the effect of communication frequency as a mediating variable. Familiar teams shared more task related information because they communicated with a higher frequency. This is in line with prior research that suggests that familiar teams are more effective at pooling information (Gruenefeld, Mannix, Williams & Neale, 1996).

Whereas *Chapters III* and *IV* focused on communication, coordinating behaviours and familiarity in multi-agency teams in crises, *Chapter V* explored cohesion in expeditions teams in ICE. The first part of the study compared pre- and post-expedition scores of cohesion in five teams, with results showing increases in cohesion in two out of the five teams. To interpret this finding, I assessed the extent to which the change in cohesion could be explained by the personality composition of the team. Results showed that extraversion had a negative effect on the development of cohesion, in line with prior research in an analogous context that found low levels of extraversion to be beneficial in ICE (Palinkas, Gunderson, Holland, Miller & Johnson, 2000).

The most important contribution of *Chapter V* was to explore day-to-day changes in cohesion over the course of the expeditions and assess to what extent this could be explained by daily events. The results of the diary data indicated that cohesion was strongly correlated with daily ratings of performance, representing the first study to use a validated measure of cohesion to show the cohesion-performance relationship in ICE. The results support the examination of how cohesion operates in ICE, given its important relationship to perceptions of performance within the team (Stuster, 2011; Vessey & Landon, 2017). The positive relationship between cohesion and performance identified here is consistent with research from a range of other contexts (Beal, Cohen, Burke & McLendon, 2003; Salas, Grossman, Hughes & Coultas, 2015).

Analyses showed a number of daily events that significantly predict day to day fluctuations in cohesion. When individuals felt satisfied with their ability to cope and experienced a sense of achievement, team cohesion increased. This is in line with prior research on conventional teams, which suggests that achieving shared goals (i.e., a sense of achievement in our study), can increase social identity in teams which fosters the development of cohesion (Salas & Cannon-Bowers, 2001). Results also demonstrated that cohesion decreased when team members were experiencing problems with their physical health, in line with previous research that found problems with physical health increased tension in teams in ICE (Sandal, Bye, Van de Vijver, 2011). In answering RQ1, findings confirm the importance of cohesion in supporting effective teamwork in ICE, directly answering the call for more research monitoring the cohesion-performance relationship over time (Greer, 2012; McClurg et al., 2017). Further, the results suggest day-to-day individual experiences influence day to day fluctuations in cohesion, illustrating they are an important consideration in maintaining cohesive teams over time.

Overall, findings across the empirical Chapters demonstrate many of the factors we have come to understand as important for effective teamwork in conventional teams are important in ETs too. Examples of findings concurrent with the conventional team literature include the relationship between: (i) communication and coordination in *Chapters III* and *IV;* (ii) familiarity and team behaviours in *Chapter IV* and (iii) cohesion and performance in *Chapter V* (Beal et al., 2003; Huckman et al., 2012; Rosen et al., 2008). There were, however, some interesting differences.

Whilst the findings on communication in *Chapters III* and *IV* are generally in line with prior research in conventional teams (i.e., communication is important), we identified one interesting difference. *Chapter IV* showed that in smaller, co-located MTS, the frequency of communications was correlated with behavioural outcome measures and played an important role in driving the familiarity-performance relationship. This contrasts to findings in the conventional team literature, which suggests only certain forms of communication (e.g., planning related, task related) are conducive to performance (Marlow et al., 2018). Whilst I do not suggest all communication, irrelevant of its content, is of equal importance in ETs, findings here emphasise the critical need to get people talking when they are responding in crisis. Waiting to find the “correct words” or to clearly formulate planning may lead to delays and reduced shared awareness amongst inter-agency partners.

Further, despite research in the conventional team literature reporting a positive association between extraversion and team cohesion (Van Vianen & DeDreu, 2001), we found that more extraverted teams were less cohesive over time in our study of teams in ICE. This is in line with prior research and is associated with extraverted individuals struggling when their social interactions are confined to a small number of individuals each day (as on expedition) (*see* Bartone, Kreuger & Barton, 2018; Palinkas et al., 2000). Over time this may increase conflict within the team and reduce overall cohesion. Interestingly, when studying MTS in crisis (*Chapter IV)*, extraversion positively correlated with two of the positive behaviours (*sharing task related information* and *sharing resources*). Whilst coordinating behaviours and cohesion are not one and the same, this finding suggests that there is a need to consider intra-extreme differences when developing models to understand and improve teamwork in extremes.

The differences highlighted between conventional teams and ETs suggest that there may be a need to generate context specific training interventions to support effective teamwork in ETs and that we cannot assume findings from the conventional team literature will transfer across. Further research is needed to continue exploring the applicability of findings in the conventional team literature to those operating in extremes and to better understand how contextual challenges can attenuate or amplify the drivers of effective teamwork.

Furthermore, the findings in relation to extraversion in MTS and teams in ICE, together with the salient challenges present in different types of extreme environment, as highlighted in *Chapter I*, suggest that future research has much to gain from clearly delineating between different types of ET. Such delineation will allow researchers to focus on variables that have the most relevance to team performance in each context (as shown throughout the empirical Chapters) and may also identify differences in how variables operate across contexts (as shown by the contrasting findings for extraversion). Further, it will ensure that findings from one type of ET can be applied more generally to other ETs with the same defining features (e.g., two different forms of MTS). This is important given the limited empirical research in this field and may expedite our understanding of teams operating in extreme and challenging settings.

RQ 2: What methods and analytical techniques are suitable for studying teams in different types of extreme environment?

Collecting data across two distinct contexts allowed me to consider the methods and analytical approaches appropriate for studying different kinds of ET. The primary methodological contribution of *Chapters II, III* and *IV* was to demonstrate the utility of immersive simulations in the study of ETs. Reviews of simulation studies exist elsewhere, however, they tend to be focused on a single ET (e.g., medical emergency teams, Cheng et al., 2016). In this thesis, I address this limitation by reviewing the use of simulation studies across a broad range of ETs (*Chapter II)*, identifying how simulation researchers can make use of recent technological and analytical advances. The use of simulation studies in researching MTS is then illustrated in subsequent empirical chapters (*Chapter III* and *IV*), in simulation design, data collection and data analysis.

A further contribution in the empirical Chapters of MTS was to show the relative utility of using practitioner and student-based samples. Practitioner-based samples are the “gold standard” when seeking to understand the experiences of ETs as they hold the necessary expertise that influences behaviours and decision-making (*see* Boulton & Cole, 2018). Specifically, in *Chapter III,* the sample of emergency responders held all the necessary knowledge relating to policies and procedures that outline guidelines for major incident response (e.g., Civil Contingencies Act, 2004; JESIP, 2013). Further the simulation was designed by training coordinators from the emergency services, ensuring that findings were more likely to represent the challenges present during incident response. This was especially useful to develop the coding dictionary of behaviours, ensuring the behaviours identified were relevant to understanding these teams and how they operate *in situ.*

However, practitioner samples tend to be small (*N* = 3 simulations in *Chapter III*), restricting the use of complex analytical techniques and challenging the generalisability of findings across contexts. Indeed, whilst improvements in team behaviours coincided with changes in communication structures in *Chapter III*, it was not possible to test for causality due to a small sample size. In student-based studies, a larger amount of data can be collected across repeat simulations, facilitating more complex statistical analysis, enabling researchers to distil how variables interact with one another and how this influences team behaviours (Aiken & Hanges, 2012). Moreover, using a practitioner-based sample can restrict the extent to which variables can be manipulated. For example, it would have been logistically impossible within the restricted time period of this thesis to source a sufficient sample of practitioners in the emergency services to manipulate familiarity across repeated simulations (e.g., *N* = 22 simulations in *Chapter IV*). This highlights how in an emerging field of research, student-based simulation studies can be used to provide "proof of concept", before seeking to test if findings transfer into an expert population. Similar benefits were obtained using a student sample in *Chapter V*. We were able to collect a larger volume of data with relative ease (*N* = 42) to trial the diary methodology in comparison to recruiting several teams of adult expedition-goers.

Data collection in *Chapters III* and *IV* illustrated how simulation studies with MTS allow researchers to capture the entirety of communications between team members. This data can be analysed to generate an understanding of how team members communicated (here, via social network analysis) and the type of behaviours employed in task completion (here, via thematic analysis to identify team behaviours). Network analysis can demonstrate how members from different component teams communicate with one another and the relative prominence of different component teams, which is helpful in advancing understanding of MTS (Shuffler, Jiménez-Rodríguez, & Kramer, 2015; Fodor & Flestea, 2016). By collecting data at multiple time points in *Chapter III*, SNA was used to show how communications changed across time and in response to changing contextual demands. SNA can also provide a frequency count of the total sum of communications between team members. For example, in *Chapter IV,* the results indicated that familiar team members communicated with a higher frequency and that this, in part, led to improved team behaviours. Thematic analysis of team member interactions can be used to generate a more complex understanding of team behaviours (Waller & Kaplan, 2018). In *Chapters III* and *IV,* I demonstrated the utility of this approach in studying MTS, by identifying coordinating behaviours and using this as a way of assessing team performance.

In studying expedition teams in ICE, the primary contribution was to demonstrate the utility of daily diaries to monitor team cohesion over time. This showed that it is possible to collect data "in the wild" to advance understanding of team dynamics in remote, isolated teams (Kozlowski, 2015). Similar methods have been used by researchers to monitor daily events, fluctuations in positive and negative affect and coping mechanisms in ICE (Smith, Kinnafick & Saunders, 2017; Smith, Barrett & Sandal, 2018). An important advancement here was to evidence how the methods can be used to capture perceptions of cohesion and relate this to changes in daily events. I further demonstrated how studying expedition teams offers an attractive analogous setting to understanding other, harder to reach teams in ICE. Research has demonstrated how findings from expedition teams can apply to teams involved in long distance spaceflight (Sandal, Leon & Palinkas, 2006), and there are recent suggestions this is also true for special forces teams engaging in counter terrorism operations (Smith & Barrett, 2019).

Studying expedition teams also highlighted the use of several analytical techniques in studying ETs. The diary data were analysed using multi-level modelling, which is especially useful in accounting for the individual variance in participants' responses over time and produces reliable results even with small sample sizes (Bell, Morgan, Schoenberger, Kromrey & Ferrron, 2014). In addition, Bayesian statistics were used in conjunction with classical significance testing when comparing pre- and post-expedition scores of cohesion. Prior research shows that only marginal improvements in cohesion are observed when comparing scores over time (Chang & Bordia, 2001). This is likely to do with team members wishing to give a favoured impression of their team. A Bayesian approach enables researchers to account for this by comparing the statistical model of the data from the study (in this case, cohesion over time in the expedition teams), against the posterior probability distribution (prior knowledge attained either from previous research, subject matter experts or pilot work) (Zyphur & Oswald, 2015). This can identify differences across groups, even if sample sizes are small and differences between scores minor. Whilst a significant overall change in cohesion was identified using classic significance testing in this instance, the size of Bayes factor does more to demonstrate the magnitude of this difference, by accounting for findings in previous research. This shows the promise of Bayes in the study of ETs, with other uses of Bayes highlighted in *Chapter II,* when discussing emerging analytical techniques in simulation research.

Overall, in this thesis I demonstrate why researchers must seek to employ methods and analytical techniques that are suited to the context in which ETs operate.

The choice of which method and analytical technique employed not only relates to availability and feasibility (e.g., it is challenging to use simulations on expedition teams as they operate for long periods in physically demanding environments), but also a careful consideration of the questions that need answering. When teamworking episodes are likely to last hours (e.g., MTS strategic command in the response to a major incident), monitoring real-time changes in behaviour as it emerges minute by minute will have relevance to understanding teamwork. Simulations offer an opportune platform with which to do this in a controlled and tightly monitored setting. However, when teamworking episodes last weeks, perhaps even months (e.g., expedition teams in ICE), daily measurements of variables are likely to be sufficient in generating a greater understanding of teamwork. Diary methods are suited to this, allowing researchers to carefully monitor micro-variations in the day-to-day working of teams who must work and live alongside one another for prolonged periods. Accordingly, there ought to be a degree of proportionality in the type and format of data collected dependent on whether teams are fast-forming and quickly disbanding (MTS) or if they are long-lasting and emerge over time (ICE).

The examples given in this thesis provide direction for future study in this field and it is hoped this encourages the much-needed empirical research, in place of the review articles that exist in abundance. This will enable ET researchers to move beyond theoretical frameworks and towards a clearer understanding of what factors support teamwork in stressful, high-stakes settings. Doing so would enable researchers to provide empirically informed solutions to equip teams operating in extremes.

Practical implications

From a theoretical perspective, studying ETs is important to ascertain how models and frameworks developed with conventional teams apply to those working in extreme and challenging settings. Given the frequency with which errors in extremes are attributed to teamwork failures, the practical implications of this thesis are of equal importance. The overall conclusions of thesis in relation to context highlight how leaders of ETs may benefit from considering the contextual factors that confine and restrain how team members interact. For instance, the findings in relation to communication in MTS would suggest the size of the team and whether or not teams are geographically dispersed effects the importance of network centralisation. That is, when teams are larger and more disparate, having a highly centralised network can reduce the effectiveness of team behaviours and reduce shared awareness. In such instances, practitioners should consider implementing decentralised structures to ensure information is shared effectively across multi-agency partners. When teams are smaller and co-located, the relative distribution of communications across the group (e.g., network centralisation) may not matter as much as each team member is present and able to hear all information that is being shared at any one time. Here, practitioners should focus on ensuring that all key information is shared, as it may not matter as much that communications are distributed evenly across agencies.

The empirical Chapters on MTS further illustrate how communication is vital to improving team behaviours in emergency response (*see* Fodor & Flestea, 2016; Waring et al., 2019). This has important implications for future training and echoes many of the recent reports into the continued challenges of multi-agency emergency response teams. In a recent U.K report following the Manchester Arena Attack, an inquiry found that "almost every organisation found that improvements needed to be made in its ability to communicate within the organisation and externally" and many of the challenges during the response were attributed to poor inter-agency communication (Kerslake, 2018, pg. 207). Findings in *Chapter IV* suggest increasing inter-agency training (and therefore increasing familiarity) might be one way to alleviate this. However, given that the effect of familiarity was, in part, due to an increase in overall communications in familiar teams, similar improvements might be found by simply encouraging teams to communicate more openly with inter-agency colleagues.

A further important implication of the findings is that centralised communications and decision making may be damaging in the response to emergencies. In our study of commanders in the emergency services (*Chapter III*), we found that networks remained centralised throughout, albeit less so in the latter phases of incident response. Centralised networks work well in scenarios which are well structured and in which there are designated "boundary spanners": individuals within the MTS whose role it is to maintain communication channels and coordinated action across component teams (Firth, Hollenbeck, Mile, Ilgen & Barnes, 2015; Leavitt, 1951). However, often in emergency response, the individual with the over-arching responsibility to make key decisions is also tasked with maintaining effective communications across agencies (Waring, Moran & Page, 2020). This was evident in our simulation scenario as it was the Chair who was driving the decision making and directing the majority of communications across the system.

The over-reliance on a single individual to manage communications and decision making creates challenges and this can again be exemplified in the review of Manchester Arena Attack. Whilst the Force Duty Officer was praised for their dynamic decision making, it was said that the demands place on them directly contributed to "significant difficulties in getting through to talk to him" (Kerslake, 2018, pg. 127).

This contributed to failures in communicating effectively with multi-agency partners, in particular with the Fire Service who were delayed by two hours in their response. Visualising the network structures in *Chapter III* demonstrated the potential importance of a “boundary spanning role” in the response to emergencies, and such a role may have reduced or eliminated some of the failures identified by Kerslake. Future research should explore the effects of implementing a boundary spanning role in the response to emergencies. It is likely this would lessen the load on the commanding officer and improve shared awareness across agencies during critical phases of the response.

The coordinating behaviours I identified also provide direction for targeted training to improve coordination in MTS. For example, findings identified *role uncertainties* as disruptive to team coordination, suggesting that improving understanding of the roles and responsibilities of different component teams will improve teamwork in MTS. Prior research has shown that role clarification training can improve team performance in medical teams and similar benefits may be observed in emergency responders (Klein, DiazGranados, Salas Burke, Lyons & Goodwin, 2009; Salas, DiazGranados, Weaver, & King, 2008). As an alternative to role clarification training, a designated boundary spanner might be used to establish a shared sense of understanding of roles and responsibilities across inter-agency partners, ensuring team leaders are aware of how decisions might conflict/compromise those of other agencies.

The findings in *Chapter V* show the importance of cohesion for team performance in ICE and demonstrate how fluctuations in cohesion are predicted by day-to-day variations in team members experiences. These findings have two important implications for practitioners. First, they suggest there is potential to use cohesion as a way of monitoring effective teamwork in ICE. One of the central challenges for teams in ICE is how to ensure teams can perform effectively, and in close contact with one another for prolonged periods. Maintaining cohesion offers a solution to this challenge as it encapsulates both the social and task-oriented aspects of teamworking (Vessey & Landon, 2017). Further, findings suggest team leaders may mitigate breakdowns in cohesion by monitoring fluctuations in day-to-day events. For example, if certain team members are experiencing problems with their physical health, leaders should be aware that this could have consequences for the dynamics of the entire team. By understanding the role of daily events in the development of cohesion, it might be possible to intervene/offer support for team members experiencing difficulties, before this affects the wider group.

Limitations and future directions

Despite important theoretical and practical implications, this thesis is not without limitations. A detailed discussion of limitations can be found in each Chapter; however, it is important to acknowledge those that span this body of work as a whole. In acknowledging the limitations to this thesis, I provide some further considerations for future avenues of research.

Scope and framing. The first limitation relates to the scope of findings in this thesis. As referenced in *Chapter I,* teamwork is a multi-faceted phenomenon with many factors at play that may either contribute to, or disrupt performance (Rosseau, Aube & Savoie, 2006). Each aspect of teamwork including team composition, team processes, emergent states and contextual factors are interrelated, in which a change in one may result in a change in the other (Dinh & Salas, 2017). The factors studied in this thesis are shown to be important to support effective teamwork in ETs, however they may emerge to be less important when we consider other factors. For example, personality explained 20% of the variance in pre- and post-expedition cohesion scores, however, we may find that this variance reduces if we include other factors such as the leadership style of each team.

However, this is a problem that can be attributed to all team research and indeed psychology research in general (Salas, Reyes & Daniel, 2018). Findings are not intended to be exhaustive (determining all the factors that support teamwork in ETs); rather I seek to extend prior research by considering the relative challenges present in different types of extreme environment and using this to direct empirical research. By addressing some of the problems with prior research (e.g., failing to take measurements over time, failing to use validated measures of cohesion), findings still offer an important contribution to the limited empirical evidence of how teamwork operates in extremes.

A second limitation is the lack of multi-level measures of teamwork. Teamwork is a multi-level phenomenon which emerges and changes over time and is influenced by factors residing at the individual, team, and system level (Kozlowski & Chao, 2018). I acknowledge the importance of collecting data from multiple levels in *Chapter II* when discussing the use of simulations in the study of ETs, and in *Chapters III* and *V* by measuring teamwork over time, however, notable omissions remain. In studying MTS, all measures were taken at the system level, driven by an over-arching aim to understand inter-team processes. However, this approach fails to account for how intra-team behaviours and component team performance impacts the MTS as a whole. It is likely that the factors contributing to component team performance will be different to those that contribute to system level performance (DeChurch & Zaccaro, 2013). Further, despite the recognised importance of individual behaviours such as leadership and non-technical skills for ETs, the studies of MTS did not attend closely to teamwork behaviours at the individual level (*see* Carter, DeChurch & Zaccaro, 2014; Murase, Carter, DeChurch & Marks, 2014; Yule, Flin, Paterson-Brown, Maran & Rowley, 2006). In contrast to the MTS studies, the study of expedition teams collected all measures from the individual level (e.g., individual perceptions of team cohesion). This is usual in the ICE literature and served its purpose in understanding how individual perceptions of the team were influenced by day-to-day variations in experiences. However, aggregating data at the individual level to infer understanding about team-based variables can create challenges (Golden, Chang & Kozlowski, 2018). For instance, there may have been differences in how individuals perceived team cohesion and how cohesive a team actually behaved over the course of the day.

Generating a more complex understanding of ETs will require future empirical research to account for variance at the individual, team and, for MTS, the system level. This will likely represent the next frontier in team research and lead to the development of complex, multi-level, non-linear models of effective teamwork. Such studies may be more feasible in conventional teams (where participants may be easier to access, and larger sample sizes achieved); however, these are critical issues for ETs too. The advent of wearable technology and its use in simulation studies (as outlined in *Chapter II)* will likely hold the key to building the empirical basis for developing more sophisticated models of MTS teamwork. Wearable devices could track changes in within-person behaviours (e.g., measuring emerging leadership on the basis of speaking frequency, Chaffin et al., 2017), in addition to team processes at the component and system level, with relative ease in comparison to the labour-intensive task of recording and transcribing verbal interactions. For example, future research might compare how communication at the intra-team level affects inter-team communication and if this relationship is moderated by team leadership. In studying teams in ICE, findings may be enriched by including the team leaders' perception of team functioning as a comparison to individual perceptions. This could offer an opportunity to develop theory on how breakdowns in cohesion and or conflict may arise when leaders' perceptions are incompatible with team members. Further, recent developments may allow researchers to monitor an individual's cognitive function over time (through a reaction time game) in ICE (*see* https://www.driftextremes.com/the-tool). Future research might compare how variables at the individual level (i.e., cognitive function) relate to team-based variables such as cohesion.

A third limitation relates to how findings have been framed in terms of performance indices. In each Chapter the emphasis is on how factors (such as familiarity, or contextual demands) influence team processes and emergent states, rather than measures of overall performance. This is not uncommon in the study of ETs (Bell, Brown & Mitchell, 2019) because of the difficulties of measuring objective performance in extremes where decisions often have no right or wrong answer (Alison, Power, van den Heuvel & Waring, 2015a). However, framing the research in this way is a departure from what is often seen in the conventional team literature and may limit the generalisability of findings as processes and emergent states are not compared against a clearly defined team performance measure (*see* Espinosa et al., 2007; Mathieu, Kukenberger, D'Innocenzo & Reilly, 2015). Further it can limit the extent to which we can conclusively state if a team did or did not do well. For instance, in *Chapter III*, I focused on comparing how teams communicated and coordinated during different phases of emergency response. In *Chapter IV* I used a measure of coordinating behaviours as an outcome indicator to compare how familiar and unfamiliar teams behaved. Performance was measured in *Chapter V*, but it was measured using a single item, and the main focus of the study was instead on how cohesion emerged and was sustained in ICE.

Future research may include a greater exploration of what constitutes performance in emergency response and expedition teams. This would help to validate performance measures that can be used across studies and provide an anchor point with which to compare the relationships between different variables (*see* NOTSS, Yule et al., 2006; ANTS, Fletcher, Flin, McGeroge, Glavin, George, Maran & Patey, 2004). In the future, and as we begin to develop a greater understanding of how MTS and teams in ICE operate effectively, we may see more clearly defined performance indices in each context. This would make findings from different studies more comparable and open the door to "best practice" solutions that have been well validated across studies.

Sample size. Low sample sizes are a common challenge in the study of ETs (Bell et al., 2018). Access can be difficult for emergency response teams and responders may have limited availability to take part in research alongside their day-to-day working. For expedition teams, the amount of data that can be collected is restricted by the size of the team (teams tend to be relatively small), the limited time team members have during the day to complete surveys and the reliance on team members returning diaries at the end of their expeditions. It is therefore usual to see smaller sample sizes in the study of ETs in comparison to conventional teams (*see* Alison, Power, van den Heuvel, Humann, Palasinski & Crego, 2015b; Kahn & Leon, 1994; Mathieu et al., 2015; Smith, Barrett & Sandal, 2018; Wilkinson, Cohen-Hatton & Honey, 2019).

This defence aside, the three MTS studied from the emergency services in *Chapter III* is a relatively small number of teams from which to draw meaningful conclusions. The contribution of this Chapter can, however, be measured by the richness of the data collected, rather than the volume of it. Participants were experts, representing up to 11 agencies operating at the highest level of command within the U.K.'s Integrated Emergency Management system (Salmon, Stanton, Jenkins &

Waller, 2011). Recordings captured all of the interactions between team members, thus representing the realities of how team members behaved, as opposed to a reliance on observers' ratings or retrospective, self-report surveys. Further, the conclusions from *Chapter III* are strengthened by the supporting evidence from a larger volume of data in the laboratory-based study in *Chapter IV.* Particularly in regard to the importance of communication in MTS and in developing the coding dictionary for MTS coordinating behaviours. In the future, it will be important to replicate findings using a larger sample of emergency responders and to see if findings apply to other, similarly fast-forming MTS in crisis. This will be important in justifying the distinction between MTS and ICE as findings from one sub-set of ET should be comparable to those with the same defining features. For instance, we would expect the coordinating behaviours identified here to apply to other MTS responding in similarly dynamic, high stakes environments (e.g., medical emergency teams, disaster response).

The sample size in *Chapter V (N* = 71) is considerably higher than those typically found in the expedition team literature (*see* Kahn & Leon, 1994; Smith et al., 2018). Further, data were collected from five teams meaning comparisons could be made across the teams to understand changes in cohesion over time. However, the generalisability of findings remains somewhat limited by the fact that this was a pilot study, designed to trial a methodology of tracking cohesion over time. I planned to include data from additional studies, in which I employed a revised version of the diary (*see* Appendix 5) to validate the method and test if findings could be replicated in an adult sample. However, in each instance, insufficient data were collected for analysis (53 diaries were distributed to five teams with less than half of these returned and completed). Data collection in expedition teams requires a great deal of compliance and commitment from participants and there are many reasons for incomplete data collection. For example, diaries might get lost on expedition or not returned following the end of the expedition. Alternatively, participants may fail to complete the diaries on a sufficient number of days or may choose to only complete certain parts of the diary.

Findings from the planned follow up diary studies are not reported in this thesis due to incomplete data collection, however, the additional measures included in the diaries offer some suggestions for future research. For instance, I included a measure of positive and negative affect, as prior research suggests that daily events can predict affect in teams and further research shows that negative affect can reduce perceptions of cohesion (Jordan, Lawrence & Troth, 2006; PANAS, Watson, Clark & Tellegen, 1988; Smith et al., 2018). The aim was to explore the relationship between individual level factors (e.g., mood), contextual changes (e.g., daily events) and team level factors (e.g., cohesion). Implementing the revised diary offers an opportunity to begin testing multi-level aspects of teamwork for teams in ICE and gain a greater understanding of how negative emotions impact group cohesion over time. Further I included open-ended questions, asking individuals to write a brief overview of what they achieved each day. Open-ended questions might be used in future research to ascertain a greater understanding of what constitutes effective performance for teams in ICE and validate quantitative measures by comparing written responses with responses given in tick box format (i.e., diary measures).

Qualitative analysis. Qualitative research offers an opportunity to gain a finer-grained and more nuanced understanding of teamwork in MTS operating in real world contexts, facilitating an empirically grounded approach to understanding how processes contribute to team effectiveness (Shuffler et al., 2015). This is especially important in instances in which we have limited knowledge (e.g., behaviours that underpin coordination in MTS, Wijnmaalen et al., 2018). However, qualitative analysis has been referred to as subjective, which may question the validity and utility of findings (Waring et al., 2018). Many recent reviews of team research make no mention of qualitative methods at all (Mathieu et al., 2018; Salas et al., 2018) and there is some suggestion that qualitative methods are a less scientific approach than quantitative techniques (Malterud, 2001).

In *Chapter III* and *Chapter IV,* transcripts of team communications were thematically analysed to identify behaviours that either disrupted or supported teamwork. This type of qualitative analysis can lead to a fixation on certain behaviours, whilst missing other equally important behaviours (Rosen & Dietz, 2017). For instance, based on my understanding of teams and prior reading of the MTS and emergency response literature, it is possible there was a degree of bias in my identification of coordinating behaviours when developing the coding dictionary. In an attempt to be as transparent as possible, I made it clear when behaviours included in the coding dictionary were based on prior research (e.g., *joint decision-making, sharing task related information).* Further, intra and inter-rater reliability was used throughout, reducing the likelihood of bias when agreement is obtained across multiple coders (LeBroten & Senter, 2008; Waller & Kaplan, 2018). I also spent some time attending training events and speaking to colleagues across the emergency services prior to data collection. This helped to develop what is referred to as "quasi-expertise", ensuring any conclusions from the data were from a position of understanding (Pfadenhauer, 2009).

Whilst possible limitations to the findings from the qualitative analysis remain, the benefits to collecting behavioural data from rich recordings of real time interactions out-weighs the alternative of using self-report data. Future research might consider combining qualitative analysis of transcripts with quantitative analysis derived from data collected using wearable devices (e.g., sociometric badges). This may enrich findings and ensure conclusions drawn from emerging themes in the qualitative data are substantiated by quantitative analysis. For example, the frequency and intensity of speech patterns captured by wearable devices can provide an indication of how often team members are interrupted by their colleagues (Kim, McFee, Olgun, Waber & Pentland, 2012). This type data could be compared against qualitative coding of transcripts (e.g., the number of times *conflicting priorities* were identified in team interactions). This would likely strengthen our understanding of how different types of conflict impact MTS performance and generate a deeper insight into how communications are managed within teams.

Self-report measures. Whilst I advocate against the use of self-report measures in the study of ETs in *Chapter II*, when studying expedition teams in ICE (*Chapter V)*, self-report diary-based methods have proved effective (e.g., Atlis et al., 2004; Smith et al., 2018). Some researchers have argued against the proliferated use of self-report scales in psychological research because of the disparity between self-reported and actual behaviour (Dolinski, 2018). However, in instances in which teams are operating remotely (i.e., on expedition), there are limited alternatives. For example, it would be difficult to capture the dynamic changing environment of an expedition in a simulation study and observing teams *in situ* may lead to unwanted experimenter effects.

With a lack of existing alternatives to self-report measures in this context, several steps were taken to ensure findings remained as reliable as possible. Only validated scales were used in the daily diaries and pre- and post-expedition questionnaires: (i) cohesion (Mathieu et al., 2015); (ii) personality (Gosling, 2003) and (iii) daily events (Atlis et al., 2004; Smith et al., 2018). Further, measures of cohesion and daily events were collected on each day of the expedition. This is an improvement on the methods seen in many studies of teams in ICE which fail to achieve a high level of power as data are collected at longer intervals (e.g., once a week, twice a month) (Golden et al., 2018). Relatedly, as the data were analysed using multi-level modelling, any individual variation across participants responses was accounted for in the analyses as a random factor (Bell et al., 2014).

In the future we may see the development of mobile monitoring tools that are robust enough to withstand harsh temperatures in inhospitable expedition environments (*see* DRIFT; https://www.driftextremes.com). Collecting data electronically would allow the order of diary items to be randomised on each day which would increase the validity of the data by removing response biases. It would also reduce some of the aforementioned challenges to incomplete data collection as participants can upload their data as soon as they have an internet connection. Any data shared during the expedition could be analysed instantly which may provide the opportunity for researchers to test interventions to mitigate breakdowns in cohesion. This would help to generate further practical applications of this research and equip team leaders in ICE with the necessary tools to respond to and/or prevent breakdowns in cohesion.

Conclusion

This thesis explored teamwork in extreme environments. It took a context driven approach to examining factors that support effective teamwork in MTS and ICE and outlining the methods and analytical approaches suited to researching each sub-set of ET. Taken together, the findings of this thesis reach two over-arching conclusions. The first is that ETs can and should be differentiated into MTS and ICE. In highlighting this, this thesis strengthens the argument that “one size does not fit all when it comes to teamwork”, and that this applies to the differences between ETs, as well as those between ETs and conventional teams (Salas et al., 2015, p. 14). Distinguishing between types of ET will better direct empirical studies to research questions that are mindful of the salient challenges present in different types of extreme environments. It will also ensure that findings from the limited empirical research are more transferable to other teams with similar features. The second conclusion is that more work is needed to reduce the gap between applied and theoretical approaches to studying ETs. Applied studies are useful in providing practical guidance to those working in the field, but the interpretation of findings may be limited if studies fail to acknowledge theoretical advancements. Likewise, whilst theoretical papers offer frameworks with which to understand and research ETs, without empirical validation they provide no indication of how this might be of use to practitioners working in high-risk domains. By presenting a series of empirical studies that are sensitive to the applied and theoretical benefits of researching ETs, it is hoped future researchers are mindful of the need to more closely marry the two. This will be essential to advancing our understanding of teams operating in extreme environments and may reduce the gap between what is known about conventional teams and what is known about ETs.

Bibliography

Acton, B. P., Braun, M. T., & Foti, R. J. (2019). Built for unity: assessing the impact of team composition on team cohesion trajectories. *Journal of Business and Psychology*, https://doi.org/10.1007/s10869-019-09654-7

Adams, S. L., & Anantatmula, V. (2010). Social and behavioral influences on team process. *Project Management Journal*, *41*(4), 89-98.

Aiken, J. R., & Hanges, P. J. (2012). Research methodology for studying dynamic multiteam systems: Application of complexity science. In S. J. Zaccaro, M. A. Marks, & L. A. DeChurch (Eds.), *Multi-team Systems: An Organization Form for Dynamic and Complex Environments* (pp. 431-458). Routledge Taylor & Francis Group.

Alge, B. J., Wiethoff, C., & Klein, H. J. (2003). When does the medium matter? Knowledge-building experiences and opportunities in decision-making teams. *Organizational Behavior and Human Decision Processes*, *91*(1), 26-37.

Alison, L., Crego, J. (2008) *Policing major incidents: leadership and major incident management.* Willan Publishing.

Alison, L., Doran, B., Long, M. L., Power, N., & Humphrey, A. (2013). The effects of subjective time pressure and individual differences on hypotheses generation and action prioritization in police investigations. *Journal of Experimental Psychology: Applied*, *19*(1), 83-93.

Alison, L., Power, N., van den Heuvel, C., & Waring, S. (2015a). A taxonomy of endogenous and exogenous uncertainty in high-risk, high-impact contexts. *Journal of Applied Psychology*, *100*(4), 1309-1318.

Alison, L., Power, N., van den Heuvel, C., Humann, M., Palasinksi, M., & Crego, J. (2015b). Decision inertia: Deciding between least worst outcomes in emergency responses to disasters. *Journal of Occupational and Organizational Psychology*, *88*(2), 295-321.

Alison, L., van den Heuvel, C., Waring, S., Power, N., Long, A., O'Hara, T., & Crego, J. (2013). Immersive simulated learning environments for researching critical incidents: A knowledge synthesis of the literature and experiences of studying high-risk strategic decision making. *Journal of Cognitive Engineering and Decision Making*, *7(3)*, 255-272.

Allison, M. T., Duda, J. L., & Beuter, A. (1991). Group dynamics in the Himalayas. *International Review for the Sociology of Sport*, *26*(3), 175-191.

Amacher, S. A., Schumacher, C., Legeret, C., Tschan, F., Semmer, N. K., Marsch, S., & Hunziker, S. (2017). Influence of gender on the performance of cardiopulmonary rescue teams: a randomized, prospective simulator study. *Critical Care Medicine, 45*(7), 1184-1191.

Anania, E. C., Disher, T. J., Kring, J. P., Iwig, C. L., Keebler, J. R., Lazarra, E. H., & Salas, E. (2017). Communication in the spaceflight multi-team system: training and technology recommendations to support boundary spanners. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 150-154.

Anderson, C., & Talsma, A. (2011). Characterizing the structure of operating room staffing using social network analysis. *Nursing Research*, *60*(6), 378-385.

Annett, J., & Stanton, N. A. (Eds.). (2000). *Task Analysis*. CRC Press.

Asencio, R., & DeChurch, L. A. (2017). Assessing collaboration within and between teams: A multiteam systems perspective. In A. A. von Davier, M. Zhu, & P. C. Kyllonen (Eds.), *Innovative Assessment of Collaboration* (pp. 37–50). Springer.

Atlis, M. M., Leon, G. R., Sandal, G. M., & Infante, M. G. (2004). Decision processes and interactions during a two-woman traverse of Antarctica. *Environment and Behavior*, *36*(3), 402-423.

Austin, J. R. (2003). Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology*, *88*(5), 866-878.

Backlund, P., Engstrom, H., Hammar, C., Johannesson, M., & Lebram, M. (2007). Sidh-a game based firefighter training simulation. In *Information Visualization, 2007. IV'07. 11th International Conference* (pp. 899-907).

Baker, B. G., Bhalla, A., Doleman, B., Yarnold, E., Simons, S., Lund, J. N., & Williams, J. P. (2017). Simulation fails to replicate stress in trainees performing a technical procedure in the clinical environment. *Medical Teacher*, *39*(1), 53-57.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63-105.

Balkundi, P., & Harrison, D. A. (2006). Ties, leaders, and time in teams: Strong inference about network structure's effects on team viability and performance. *Academy of Management Journal*, *49*(1), 49-68.

Barrett, E., Martin, P. (2014) *Extreme: Why Some People Thrive at the Limits*. Oxford University Press.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*(3), 377-391.

Bartone, P. T., Krueger, G. P., & Bartone, J. V. (2018). Individual differences in adaptability to isolated, confined, and extreme environments. *Aerospace Medicine and Human Performance*, *89*(6), 536-546.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2012). lme4: Linear mixed-effects models using Eigen and S4 R package version 10-4 Available: http. *CRANR-projectorg/package= lme4*.

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers, 23*(1), 4-17.

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). Cohesion and performance in groups: a meta-analytic clarification of construct relations. *Journal of Applied Psychology, 88*(6), 989-1004.

Beaubien, J. M., & Baker, D. P. (2004). The use of simulation for training teamwork skills in health care: how low can you go?. *BMJ Quality & Safety*, *13*(1), 51-56.

Becker-Beck, U. (2001). Methods for diagnosing interaction strategies: An application to group interaction in conflict situations. *Small Group Research, 32*(3), 259-282.

Bell, B. A., Morgan, G. B., Schoeneberger, J. A., Kromrey, J. D., & Ferron, J. M. (2014). How low can you go? An investigation of the influence of sample size and model 624 complexity on point and interval estimates in two-level linear models. *Methodology, 10*(1), 1-11.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: a meta-analysis. *Journal of Applied Psychology*, *92*(3), 595.

Bell, S. T., & Outland, N. (2017). Team composition over time. In E. Salas, W. B. Vessey, & L. B Landon (Eds.), *Team Dynamics Over Time* (pp. 3-27). Emerald Publishing Limited.

Bell, S. T., Brown, S. G., & Mitchell, T. (2019). What We Know About Team Dynamics for Long-Distance Space Missions: A Systematic Review of Analog Research. *Frontiers in Psychology*, *10*, https://doi.org/10.3389/fpsyg.2019.00811

Bell, S. T., Brown, S. G., Colaneri, A., & Outland, N. (2018). Team composition and the ABCs of teamwork. *American Psychologist*, *73*(4), 349-362.

Bell, S. T., Fisher, D. M., Brown, S. G., & Mann, K. E. (2018). An approach for conducting actionable research with extreme teams. *Journal of Management*, *42*(5), 1-26.

Bienefeld, N., & Grote, G. (2014). Shared leadership in multiteam systems: How cockpit and cabin crews lead each other to safety. *Human Factors*, *56*(2), 270-286.

Bihari, A., & Pandia, M. K. (2015). Eigenvector centrality and its application in research professionals' relationship network. In *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)* (pp. 510-514).

Bishop, S. L. (2004). Evaluating teams in extreme environments: From issues to answers. *Aviation, Space, and Environmental Medicine*, *75*(7), 14-21.

Bishop, S. L., Grobler, L. C., & SchjØll, O. (2001). Relationship of psychological and physiological parameters during an Arctic ski expedition. *Acta Astronautica*, *49*(3-10), 261-270.

Bishop, S. L., Morphew, M. E., & Kring, J. P. (2000). Avoiding risky teams in risky environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 44*(12), 800-804.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). UCInet for Windows: Software for Social Network Analysis. Analytic Technologies.

Borucki, C. C., & Burke, M. J. (1999). An examination of service-related antecedents to retail store performance. *Journal of Organizational Behaviour*, *20*(6), 943-962.

Boulton, L., & Cole, J. (2016). Adaptive flexibility: Examining the role of expertise in the decision making of authorized firearms officers during armed confrontation. *Journal of Cognitive Engineering and Decision Making*, *10*(3), 291-308.

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing communication sequences for team training needs assessment. *Human Factors, 40*(4), 672-679.

Boyd, M., Kim, M. S., Ensari, N., & Yin, Z. (2014). Perceived motivational team climate in relation to task and social cohesion among male college athletes. *Journal of Applied Social Psychology*, *44*(2), 115-123.

Boynton, A. C., Gales, L. M., & Blackburn, R. S. (1993). Managerial search activity: The impact of perceived role uncertainty and role threat. *Journal of Management*, *19*(4), 725-747.

Bradley, B. H., Baur, J. E., Banford, C. G., & Postlethwaite, B. E. (2013). Team players and collective performance: How agreeableness affects team performance over time. *Small Group Research*, *44*(6), 680-711.

Bradley, J. H., Foltz, C. B., White, B. J., & Wise, H. (2006). The impact of goal congruence on project team performance. *The International Journal of Applied Management and Technology*, *4*(2), 219-228.

Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education*, *40*(3), 254-262.

Braun, M. T., Kozlowski, S. W., Brown, T. A., & DeShon, R. P. (2020). Exploring the Dynamic Team Cohesion–Performance and Coordination–Performance Relationships of Newly Formed Teams. *Small Group Research*. https://doi.org/10.1177/1046496420907157

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

Brown, O., Barrett, E., Power, N. (2019) Monitoring cohesion over time in expedition teams; the role of daily events and team composition. In *Proceedings of the 14th International Naturalistic Decision-Making Conference, San Francisco, U.S.A*.

Brown, O., Power, N. (2018) *Teamwork during a simulated terrorist attack.* Presented at European Association of Psychology and Law, Turku, Finland.

Brown, T. M., & Miller, C. E. (2000). Communication networks in task-performing groups: Effects of task complexity, time pressure, and interpersonal dominance. *Small Group Research*, *31*(2), 131-157.

Burke, C. S., Shuffler, M. L., & Wiese, C. W. (2018). Examining the behavioral and structural characteristics of team leadership in extreme environments. *Journal of Organizational Behaviour*, *39*(6), 716-730.

Burke, C. S., Shuffler, M. L., Heyne, K., Salas, E., & Ruark, G. A. (2014). *Improving the Performance of Military Multiteam Systems: An Empirical Investigation* (Technical Report 2014-XX). Army Research Institute for the Behavioural and Social Sciences.

Burke, C. S., Wilson, K. A., & Salas, E. (2003). Teamwork at 35,000 feet: Enhancing safety through team training. *Human Factors and Aerospace Safety, 3*(4), 287–312.

Butts, C. T. (2008). Social network analysis with sna. *Journal of Statistical Software*, *24*(6), 1-51.

Cabinet Office (2012) Revision to emergency preparedness: glossary. Civil Contingencies Act Enhancement Programme, London.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies: Implications for training requirements and strategies. In R. A. Guzzo, E. Salas, & Associates (Eds.) *Team Effectiveness and Decision Making in Organizations*, (pp 333-380). Jossey-Bass.

Carless, S. A., & De Paola, C. (2000). The measurement of cohesion in work teams. *Small Group Research*, *31*(1), 71-88.

Carter, D. R. (2014). *The Impact of Leadership Network Structure on Multiteam System Innovation* (Unpublished doctoral thesis). Georgia Institute of Technology, Atlanta.

Carter, D. R., DeChurch, L. A., & Zaccaro, S. J. (2014). Impact of leadership network structure on the creative output of multiteam systems. *Paper Proceedings of the 74th Annual Meeting of the Academy of Management (AOM).* 16874.

Casey-Campbell, M., & Martens, M. L. (2009). Sticking it all together: A critical assessment of the group cohesion–performance literature. *International Journal of Management Reviews*, *11*(2), 223-246.

Chaffin, D., Heidl, R., Hollenbeck, J. R., Howe, M., Yu, A., Voorhees, C., & Calantone, R. (2017). The promise and perils of wearable sensors in organizational research. *Organizational Research Methods*, *20*(1), 3-31.

Chang, A., & Bordia, P. (2001). A multidimensional approach to the group cohesion-group performance relationship. *Small Group Research*, *32*(4), 379-405.

Chang, K., Chang, A., & Sha, X. (2009). Work role similarity and work familiarity between members: A tripartite view of social identity towards knowledge contribution in organizations. In *PACIS 2009 Proceedings*, 57.

Chen, R., Sharman, R., Rao, H. R., & Upadhyaya, S. J. (2008). Coordination in emergency response management. *Communications of the ACM*, *51*(5), 66-73.

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. *Medical Education, 48*(7), 657-666.

Cheng, A., Kessler, D., Mackinnon, R., Chang, T. P., Nadkarni, V. M., Hunt, E. A., Duval-Arnould, J., Lin, Y., Cook, D. A., Pusic, M., Hui, J., Moher, D., Egger, M., Auerbach, M., & International Network for Simulation-based Pediatric Innovation, Research, and Education (INSPIRE) Reporting Guidelines Investigators . (2016). Reporting guidelines for health care simulation research: Extensions to the CONSORT and STROBE statements. *Advances in Simulation, 1*(25), 1-13.

Choi, S. O., & Brower, R. S. (2006). When practice matters more than government plans: A network analysis of local emergency management. *Administration & Society*, *37*(6), 651-678.

Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*(1), 5-6.

Civil Contingencies Act (2004). A short guide (revised). Retrieved from http://www.legislation.gov.uk/ukpga/2004/36/contents/enacted

Cohen-Hatton, S. R., & Honey, R. C. (2015). Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied*, *21*(4), 395-406.

Colbert, A. E., Kristof-Brown, A. L., Bradley, B. H., & Barrick, M. R. (2008). CEO transformational leadership: The role of goal importance congruence in top management teams. *Academy of Management Journal*, *51*(1), 81-96.

Crandall, B., Klein, G., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Mit Press.

Crawford, E. R., & LePine, J. A. (2013). A configural theory of team processes: Accounting for the structure of taskwork and teamwork. *Academy of Management Review*, *38*(1), 32-48

Crichton, M. T., Flin, R., & Rattray, W. A. (2000). Training decision makers–tactical decision games. *Journal of Contingencies and Crisis Management*, *8*(4), 208-217.

Cronin, M. A. (2015). Advancing the science of dynamics in groups and teams. *Organizational Psychology Review, 5*(4), 267–269.

Cronin, M. A., & Weingart, L. R. (2007). Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review*, *32*(3), 761-773.

Cronin, M. A., Weingart, L. R., & Todorova, G. (2011). Dynamics in groups: Are we there yet? *The Academy of Management Annals*, *5*(1), 571-612.

Cross, R., Rebele, R., & Grant, A. (2016). Collaborative overload. *Harvard Business Review, 94*(1), 74-79.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 64-73.

Cuijpers, M., Uitdewilligen, S., & Guenter, H. (2016). Effects of dual identification and interteam conflict on multiteam system performance. *Journal of Occupational and Organizational Psychology*, *89*(1), 141-171.

Cumin, D., Boyd, M. J., Webster, C. S., & Weller, J. M. (2013). A systematic review of simulation for multidisciplinary team training in operating rooms. *Simulation in Healthcare*, *8*(3), 171-179.

Cumin, D., Weller, J. M., Henderson, K., & Merry, A. F. (2010). Standards for simulation in anaesthesia: creating confidence in the tools. *British Journal of Anaesthesia*, *105*(1), 45-51.

Dalal, D. K., Nolan, K. P., & Gannon, L. E. (2017). Are pre-assembly shared work experiences useful for temporary-team assembly decisions? A study of Olympic ice hockey team composition. *Journal of Business and Psychology*, *32*(5), 561-574.

Davison, R. B., Hollenbeck, J. R., Barnes, C. M., Sleesman, D. J., & Ilgen, D. R. (2012). Coordinated action in multiteam systems. *Journal of Applied Psychology*, 97(4), 808–824.

De Dreu, C. K., & Weingart, L. R. (2003). A contingency theory of task conflict and performance in groups and organizational teams. *International Handbook of Organizational Teamwork and Cooperative Working*, *88*(4), 151-166.

DeChurch, L. A., & Marks, M. A. (2001). Maximizing the benefits of task conflict: The role of conflict management. *International Journal of Conflict Management*, *12*(1), 4-22.

DeChurch, L. A., & Marks, M. A. (2006). Leadership in multiteam systems. *Journal of Applied Psychology*, *91*(2), 311-329.

DeChurch, L. A., & Mathieu, J. E. (2009). Thinking in terms of multiteam systems. In E. Salas, G. F. Goodwin, & S. Burke (Eds.) *Team Effectiveness in Complex Organizations: Cross-disciplinary Perspectives and Approaches*, (pp. 267-292). Routledge Taylor & Francis Group.

DeChurch, L. A., & Zaccaro, S. J. (2013). Innovation in scientific multiteam systems: Confluent and countervailing forces. *National Academy of Sciences Committee on Team Science.* National Academy of Sciences.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, *323*(5910), 66-69.

Dekker, A. H. (2000, October). Social network analysis in military headquarters using CAVALIER. In *Proceedings of Fifth International Command and Control Research and Technology Symposium* (pp. 24-26).

DeShon, R. P., Kozlowski, S. W., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. *Journal of Applied Psychology*, *89*(6), 1035-1056.

Dieckmann, P., Gaba, D., & Rall, M. (2007). Deepening the theoretical foundations of patient simulation as social practice. *Simulation in Healthcare*, *2*(3), 183-193.

Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review*, *25*(1), 207-218.

Dierdorff, E. C., Rubin, R. S., & Morgeson, F. P. (2009). The milieu of managerial work: an integrative framework linking work context to role requirements. *Journal of Applied Psychology*, *94*(4), 972-988.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*(1), 417-440.

Dinh, J. V., &Salas, E. (2017). Factors that influence teamwork. In Salas, E., Rico, R. & Passmore, J. (Eds.), *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*, (pp. 13-41). Wiley Blackwell.

Dionisio, J. D. N., & Gilbert, R. (2013). 3D Virtual worlds and the metaverse: Current status and future possibilities. *ACM Computing Surveys*, *45*(3), 34.

DiRosa, G. (2013). *Emergent Phenomena in Multiteam Systems: An Examination of Between-Team Cohesion.* (Unpublished doctoral thesis), George Mason University, Virginia.

Doliński, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, *13*(2), 1-14.

Driskell, J. E., & Mullen, B. (2005). Social Network Analysis. In N. A. Stanton (Eds.), *Handbook of Human Factors and Ergonomics Methods*, (pp 58.1 – 58.6). CRC.

Driskell, J. E., Salas, E., & Johnston, J. (1999). Does stress lead to a loss of team perspective? *Group Dynamics: Theory, Research, and Practice*, *3*(4), 291-302.

Driskell, T., Salas, E., & Driskell, J. E. (2018). Teams in extreme environments: Alterations in team development and teamwork. *Human Resource Management Review*, *28*(4), 434-449.

Dwyer, T., & Raftery, A. E. (1991). Industrial accidents are produced by social relations of work: A sociological theory of industrial accidents. *Applied Ergonomics, 22*(3), 167-178.

Dyer, J. L. (1984). Team research and team training: A state-of-the-art review. In F. A. Muchker, (Ed.), *Human Factors Review*, (pp. 285-323). Human Factors Society.

Earley, P. C., & Gibson, C. B. (1998). Taking stock in our progress on individualism-collectivism: 100 years of solidarity and community. *Journal of Management*, *24*(3), 265-304.

Eby, L. T., & Dobbins, G. H. (1997). Collectivistic orientation in teams: an individual and group-level analysis. *Journal of Organizational Behaviour*, *18*(3), 275-295.

Edmondson, A. (1999). Psychological safety and learning behaviour in work teams. *Administrative Science Quarterly*, *44*(2), 350-383.

Espinosa, J. A., Lerch, J., & Kraut, R. (2004). Explicit vs. implicit coordination mechanisms and task dependencies: One size does not fit all. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance.* (pp. 107-129). APA Books.

Espinosa, J. A., Slaughter, S. A., Kraut, R. E., & Herbsleb, J. D. (2007). Familiarity, complexity, and team performance in geographically distributed software development. *Organization Science*, *18*(4), 613-630.

Evans, C. R., & Dion, K. L. (2012). Group cohesion and performance: A meta-analysis. *Small Group Research*, *43*(6), 690-701.

Everitt, B. (1996). *Making Sense of Statistics in Psychology: A Second-Level Course.* Oxford University Press.

Eyre, M., Crego, J., & Alison, L. (2008). Electronic debriefs and simulations as descriptive methods for defining the critical incident landscape. In L. Alison & J. Crego (Eds.), *Policing critical incidents* (pp. 24–53). Willan.

Eys, M. A., Ritchie, S., Little, J., Slade, H., & Oddson, B. (2008). Leadership status congruency and cohesion in outdoor expedition groups. *Journal of Experiential Education*, *31*(1), 78-94.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80-92.

Festinger, L., & Thibaut, J. (1951). Interpersonal communication in small groups. *The Journal of Abnormal and Social Psychology*, *46*(1), 92-99.

Firth, B. M., Hollenbeck, J. R., Miles, J. E., Ilgen, D. R., & Barnes, C. M. (2015). Same page, different books: Extending representational gaps theory to enhance performance in multiteam systems. *Academy of Management Journal*, *58*(3), 813-835.

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space, and Environmental Medicine, 78*(5), 86-95.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., & Patey, R. (2004). Rating non-technical skills: developing a behavioural marker system for use in anaesthesia. *Cognition, Technology & Work*, *6*(3), 165-171.

Fodor, O. C., & Flestea, A. M. (2016). When fluid structures fail: A social network approach to multi-team systems' effectiveness. *Team Performance Management*, *22*(3/4), 156-180.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215-239.

Gamble, K. R., Vettel, J. M., Patton, D. J., Eddy, M. D., Davis, F. C., Garcia, J. O., Spangler, D. P., Thayer, J. F. & Brooks, J. R. (2018). Different profiles of decision making and physiology under varying levels of stress in trained military personnel. *International Journal of Psychophysiology*, *131*, 73-80.

Geersbro, J., & Ritter, T. (2010). External performance barriers in business networks: uncertainty, ambiguity, and conflict. *Journal of Business & Industrial Marketing*, *25*(3), 196-201.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, *24*(6), 997-1016.

Gerber, D. E., Reimer, T., Williams, E. L., Gill, M., Loudat Priddy, L., Bergestuen, D., Schiller, J, H., Kirkpatrick, H., Simon, J. & Craddock Lee, S. J. (2016). Resolving rivalries and realigning goals: Challenges of clinical and research multiteam systems. *Journal of Oncology Practice*, *12*(11), 1020-1028.

Gibson, C. B., & Cohen, S. G. (2003). *Virtual Teams that Work: Creating conditions for Virtual Team Effectiveness*. Jossey-Bass.

Gillespie, B. M., Gwinner, K., Chaboyer, W., & Fairweather, N. (2013). Team communications in surgery–creating a culture of safety. *Journal of Interprofessional Care*, *27*(5), 387-393.

Glew, D. J. (2009). Personal values and performance in teams: An individual and team-level analysis. *Small Group Research, 40*(6), 670–693.

Golden, S. J., Chang, C. H., & Kozlowski, S. W. (2018). Teams in isolated, confined, and extreme (ICE) environments: Review and integration. *Journal of Organizational Behaviour*, *39*(6), 701-715.

Goodwin, G. F., Blacksmith, N., & Coats, M. R. (2018). The science of teams in the military: Contributions from over 60 years of research. *American Psychologist*, *73*(4), 322-333.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Graziano, W. G., Hair, E. C., & Finch, J. F. (1997). Competitiveness mediates the link between personality and group performance. *Journal of Personality and Social Psychology*, *73*(6), 1394-1408.

Greer, L. L. (2012). Group cohesion: Then and now. *Small Group Research*, *43*(6), 655-661.

Grote, G., Kolbe, M., Zala-Mezö, E., Bienefeld-Seall, N., & Künzle, B. (2010). Adaptive coordination and heedfulness make better cockpit crews. *Ergonomics*, *53*(2), 211-228.

Gruenfeld, D. H., Mannix, E. A., Williams, K. Y., & Neale, M. A. (1996). Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational Behavior and Human Decision Processes*, *67*(1), 1-15.

Grund, T. U. (2016). The relational value of network experience in teams: Evidence from the English Premier League. *American Behavioral Scientist*, *60*(10), 1260-1280.

Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of level of analysis and task interdependence. *Small Group Research*, *26*(4), 497-520.

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*(1), 307-338.

Hackman, J. R. (1986). The psychology of self-management in organizations. In M. S. Pallak & R. O. Perloff (Eds.), *Psychology and Work: Productivity, Change and Employment*. American Psychological Association.

Hackman, J. R. (1987). The design of work teams. In J. W, Lorsch (Ed.), *Handbook of Organizational Behavior*, Prentice-Hall.

Hackman, J. R. (2002). Why teams don't work. *Theory and Research on Small Groups*. Springer.

Hackman, J. R. (2003). Learning more by crossing levels: Evidence from airplanes, hospitals, and orchestras. *Journal of Organizational Behaviour*, *24*(8), 905-922.

Haddow, G. D., & Bullock J, A. (2003) *Introduction to Emergency Management*. Butterworth Heinemann.

Hällgren, M., Rouleau, L., & De Rond, M. (2018). A matter of life or death: How extreme context research matters for management and organization studies. *Academy of Management Annals*, *12*(1), 111-153.

Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T., & Vanderstoep, S. W. (2003). Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality. *Personnel Psychology*, *56*(3), 633-669.

Healey, J., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6*(2), 156-166.

Helsloot, I. (2005). Bordering on reality: Findings on the bonfire crisis management simulation. *Journal of Contingencies and Crisis Management*, *13*(4), 159-169.

Hinsz, V. B., & Betts, K. R. (2012). Conflict in multiteam situations. In S. J. Zaccaro, M. A. Marks, L. DeChurch (Eds.) *Multiteam systems: An Organization Form for Dynamic and Complex Environments*, (pp. 289-322). Routledge Taylor & Francis Group.

Hinz, A., Brähler, E., Schmidt, P., & Albani, C. (2005). Investigating the circumplex structure of the Portrait Values Questionnaire (PVQ). *Journal of Individual Differences*, *26*(4), 185-193.

Hochmitz, I., & Yuviler-Gavish, N. (2011). Physical fidelity versus cognitive fidelity training in procedural skills acquisition. *Human Factors, 53*(5), 489-501.

Hoegl, M., Weinkauf, K., & Gemuenden, H. G. (2004). Interteam coordination, project commitment, and teamwork in multiteam R&D projects: A longitudinal study. *Organization Science*, *15*(1), 38-55.

Houghton, R. J., Baber, C., McMaster, R., Stanton, N. A., Salmon, P., Stewart, R., & Walker, G. (2006). Command and control in emergency services operations: a social network analysis. *Ergonomics*, *49*(12-13), 1204-1225.

Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science*, *55*(1), 85-100.

Huckman, R. S., Staats, B. R., & Upton, D. M. (2012). Team familiarity, role experience, and performance: evidence from Indian software services. *IEEE Engineering Management Review*, *40*(1), 99-118.

Hughes, A. M., Gregory, M. E., Joseph, D. L., Sonesh, S. C., Marlow, S. L., Lacerenza, C. N., Benishek, L. E., King, H. B. & Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare. *Journal of Applied Psychology*, *101*(9), 1266-1304.

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist*, *54*(2), 129-139.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology, 56*, 517-543.

Jaca, C., Viles, E., Tanco, M., Mateo, R., & Santos, J. (2013). Teamwork effectiveness factors in healthcare and manufacturing industries. *Team Performance Management: An International Journal*, *19*(3/4), 222-236.

James, E. H., & Wooten, L. P. (2010). *Leading under pressure: From surviving to thriving before, during, and after a crisis*. Routledge Taylor & Francis Group.

Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.

Jehn, K. A., & Shah, P. P. (1997). Interpersonal relationships and task performance: An examination of mediation processes in friendship and acquaintance groups. *Journal of Personality and Social Psychology*, *72*(4), 775-790.

JESIP (2013). Joint Doctrine: The Interoperability Framework. London: JESIP.

Jiang, Y., O'Neal, E., Rahimian, P., Yon, J. P., Plumert, J. M., & Kearney, J. K. (2016). Action coordination with agents: crossing roads with a computer-generated character in a virtual environment. In *Proceedings of the ACM Symposium on Applied Perception* (pp. 57-64). ACM.

Johns, G. (2006). The essential impact of context on organizational behaviour. *Academy of Management Review*, *31*(2), 386-408.

Johnson, C., Anderson, S., Dallimore, J., Winser, S., & Warrell, D. A. (2008). *Oxford Handbook of Expedition and Wilderness Medicine*. Oxford University Press.

Jordan, P. J., Lawrence, S. A., & Troth, A. C. (2006). The impact of negative mood on team performance. *Journal of Management & Organization*, *12*(2), 131-145.

Joshi, K., Hernandez, J., Martinez, J., Abdel-Fattah, K., & Gardner, A. K. (2018). Should they stay or should they go now? Exploring the impact of team familiarity on interprofessional team training outcomes. *The American Journal of Surgery*, *215*(2), 243-249.

Junglas, I., & Ives, B. (2007). Recovering IT in a disaster: Lessons from Hurricane Katrina. *MIS Quarterly Executive*, *6*(1), 39-51.

Kabanoff, B., Waldersee, R., & Cohen, M. (1995). Espoused values and organizational change themes. *Academy of Management Journal*, *38*(4), 1075-1104.

Kadiri, Z. O., Nden, T., Avre, G. K., Oladipo, T. O., Edom, A., Samuel, P. O., & Ananso, G. N. (2014). Causes and Effects of Accidents on Construction Sites (A Case Study of Some Selected Construction Firms in Abuja, FCT Nigeria). *IOSR Journal of Mechanical and Civil Engineering*, *11*(5), 66-72.

Kahn, P. M., & Leon, G. R. (1994). Group climate and individual functioning in an all-women Antarctic expedition team. *Environment and Behaviour*, *26*(5), 669-697.

Kapucu, N. (2006). Interagency communication networks during emergencies: Boundary spanners in multiagency coordination. *The American Review of Public Administration*, *36*(2), 207-225.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly*, *27*(1), 81-104.

Kauffeld, S., & Meyers, R. A. (2009). Complaint and solution-oriented circles: Interaction patterns in work group discussions. *European Journal of Work and Organizational Psychology*, *18*(3), 267-294.

Kayes, D. C. (2004). The 1996 Mount Everest climbing disaster: The breakdown of learning in teams. *Human Relations*, *57*(10), 1263-1284.

Kealey, D. J. (2004). Research on intercultural effectiveness and its relevance to multicultural crews in space. *Aviation, Space, and Environmental Medicine*, *75*(7), 58-64.

Kelley, K. (2013). Effect size and sample size planning. In Little, T.D. (Ed.), *Oxford Handbook of Quantitative Methods* (pp. 206–222). Oxford University Press.

Kerslake Report (2018). Retrieved from https://www.kerslakearenareview.co.uk/media/1022/kerslake_arena_review_printed_final.pdf

Keyton, J., Ford, D. J., & Smith, F. L. (2012). Communication, collaboration, and identification as facilitators and constraints of multiteam systems. In S. J. Zaccaro, M. A. Marks, & L. A. DeChurch (Eds.), *Multi-team Systems: An Organization Form for Dynamic and Complex Environments* (pp. 173–190). Routledge Taylor & Francis Group.

Killumets, E., D'Innocenzo, L., Maynard, M. T., & Mathieu, J. E. (2015). A multilevel examination of the impact of team interpersonal processes. *Small Group Research*, *46*(2), 227-259.

Kim, T., McFee, E., Olguin, D. O., Waber, B., & Pentland, A. S. (2012). Sociometric badges: Using sensor technology to capture new forms of collaboration. *Journal of Organizational Behavior*, *33*(3), 412-427.

Kinateder, M., Ronchi, E., Gromer, D., Müller, M., Jost, M., Nehfischer, M., Muhleberger, A. & Pauli, P. (2014). Social influence on route choice in a virtual reality tunnel fire. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26(A)*, 116-125.

Kirkman, B. L., & Shapiro, D. L. (2001). The impact of cultural values on job satisfaction and organizational commitment in self-managing work teams: The mediating role of employee resistance. *Academy of Management Journal*, *44*(3), 557-569.

Kjærgaard, A., Leon, G. R., & Fink, B. A. (2015). Personal challenges, communication processes, and team effectiveness in military special patrol teams operating in a polar environment. *Environment and Behavior*, *47*(6), 644-666.

Klein, C., DiazGranados, D., Salas, E., Le, H., Burke, C. S., Lyons, R., & Goodwin, G. F. (2009). Does team building work? *Small Group Research*, *40*(2), 181-222.

Klein, G. A., & Woods, D. D. (1993). Conclusions: Decision making in action. In G.A. Klein, J. Orasanu, R. Calderwood & C.E. Zsambok (Eds.) *Decision Making in Action: Models and Methods*, (pp. 404-411). Ablex.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, *18*(3), 81-85.

Klein, K. J., & Kozlowski, S. W. (2000). *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*. Jossey-Bass.

Klein, K. J., Lim, B. C., Saltz, J. L., & Mayer, D. M. (2004). How do they get there? An examination of the antecedents of centrality in team networks. *Academy of Management Journal*, *47*(6), 952-963.

Klein, K. J., Ziegert, J. C., Knight, A. P., & Xiao, Y. (2006). Dynamic delegation: Shared, hierarchical, and deindividualized leadership in extreme action teams. *Administrative Science Quarterly*, *51*(4), 590-621.

Klein, R. L., Bigley, G. A., & Roberts, K. H. (1995). Organizational culture in high reliability organizations: An extension. *Human Relations*, *48*(7), 771-793.

Knoke, D., & Yang, S. (2008). *Social Network Analysis*. Sage.

Kozlowski, S. W. (2015). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review*, *5*(4), 270-299.

Kozlowski, S. W. (2018). Enhancing the effectiveness of work groups and teams: a reflection. *Perspectives on Psychological Science*, *13*(2), 205-212.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In Kozlowski (Eds.), *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions* (pp. 3-90). Jossey-Bass.

Kozlowski, S. W. J., Biswas, S., & Chang, C.-H. (2014). *Monitoring and Regulating Teamwork: Final report* (No. NNX12AR15G). National Aeronautics and Space Administration.

Kozlowski, S. W. J., Gully, S. M., Nason, E. R., & Smith, E. M. (1999). Developing adaptive teams: A theory of compilation and performance across levels and time. In D. R. Ilgen & E. D. Pulakos (Eds.), *The Changing Nature of Work performance: Implications for Staffing, Personnel Actions, and Development*, (pp. 240-292). Jossey-Bass.

Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations. In I. E. Weiner (Eds.), *Handbook of Psychology*, (pp. 333-375). Wiley.

Kozlowski, S. W., & Chao, G. T. (2018). Unpacking team process dynamics and emergent phenomena: Challenges, conceptual advances, and innovative methods. *American Psychologist*, *73*(4), 576-592.

Kozlowski, S. W., & DeShon, R. P. (2004). A Psychological Fidelity approach to Simulation-based Training: Theory, Research and Principles. In E. Salas, L. R. Elliott, S. G. Schflett, & M. D. Coovert (Eds.), *Scaled Worlds: Development, Validation, and Applications* (pp. 75-99). Ashgate Publishing.

Kozlowski, S. W., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*(3), 77-124.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178-206.

Lacerenza, C. N., Rico, R., Salas, E., & Shuffler, M. L. (2014). The future of multiteam systems: Implications for research and practice. In In M. L. Shuffler, R. Rico & E. Salas (Eds.) *Pushing the Boundaries: Multiteam systems in Research and Practice* (pp. 223-240). Emerald Group Publishing Limited.

Lanaj, K., Hollenbeck, J. R., Ilgen, D. R., Barnes, C. M., & Harmon, S. J. (2013). The double-edged sword of decentralized planning in multiteam systems. *Academy of Management Journal*, *56*(3), 735-757.

Landon, L. B., Slack, K. J., & Barrett, J. D. (2018). Teamwork and collaboration in long-duration space missions: Going to extremes. *American Psychologist*, *73*(4), 563-575.

Landon, L. B., Vessey, W. B., & Barrett, J. D. 2015. *Risk of performance and behavioral health decrements due to inadequate cooperation, coordination, communication, and psychosocial adaptation within a team*. Technical report, NASA Johnson Space Center, Houston.

Larsen, S., Nystad, E., & Taylor, C. (2014). Coordinating mechanisms for multiteam systems: A case study from the petroleum industry. In M. L. Shuffler, R. Rico & E.

Salas (Eds.) *Pushing the Boundaries: Multiteam Systems in Research and Practice* (pp. 115-133). Emerald Group Publishing Limited.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, *46*(1), 38-50.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815-852.

Leenders, R. T. A., Contractor, N. S., & DeChurch, L. A. (2016). Once upon a time: Understanding team processes as relational event networks. *Organizational Psychology Review, 6*(1), 92-115.

Leon, G. R. (1991). Individual and group process characteristics of polar expedition teams. *Environment and Behavior*, *23*(6), 723-748.

Leon, G. R., Kanfer, R., Hoffman, R. G., & Dupre, L. (1994). Group processes and task effectiveness in a Soviet-American expedition team. *Environment and Behaviour*, *26*(2), 149-165.

Leon, G. R., Sandal, G. M., Fink, B. A., & Ciofani, P. (2011). Positive experiences and personal growth in a two-man North Pole expedition team. *Environment and Behaviour*, *43*(5), 710-731.

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, *61*(2), 273-307.

Lester, E., Georgiou, A., Hein, M., Littlepage, G., Moffett III, R., & Craig, P. (2017). Improving Aviation Students' Teamwork, Problem Solving, Coordination, and Communications Skills During a High-Fidelity Simulation. In *19th International Symposium on Aviation Psychology* (p. 119).

Littlepage, G., Robison, W., & Reddington, K. (1997). Effects of task experience and group experience on group performance, member ability, and recognition of expertise. *Organizational Behaviour and Human Decision Processes*, *69*(2), 133-147.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705-717.

Luciano, M. M., DeChurch, L. A., & Mathieu, J. E. (2018). Multiteam systems: A structural framework and meso-theory of system functioning. *Journal of Management*, *44*(3), 1065-1096.

Lynch, S.M. (2010). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer.

Mallaro, S., Rahimian, P., O'Neal, E. E., Plumert, J. M., & Kearney, J. K. (2017). A comparison of head-mounted displays vs. large-screen displays for an interactive pedestrian simulator. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (pp. 1-4).

Malterud, K. (2001). Qualitative research: standards, challenges, and guidelines. *The Lancet*, *358*(9280), 483-488.

Manser, T., Dieckmann, P., Wehner, T., & Rall, M. (2007). Comparison of anaesthetists' activity patterns in the operating room and during simulation. *Ergonomics*, *50*(2), 246-260.

Manzey, D., & Lorenz, B. (1998). Mental performance during short-term and long-term spaceflight. *Brain Research Reviews*, *28*(1-2), 215-221.

Maran, N. J., & Glavin, R. J. (2003). Low-to high-fidelity simulation–a continuum of medical education? *Medical Education*, *37*(1), 22-28.

Marks, M. A. (2000). A critical analysis of computer simulations for conducting team research. *Small Group Research*, *31*(6), 653-675.

Marks, M. A., DeChurch, L. A., Mathieu, J. E., Panzer, F. J., & Alonso, A. (2005). Teamwork in multiteam systems. *Journal of Applied Psychology*, *90*(5), 964-971.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*(3), 356-376.

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology*, *85*(6), 971-986.

Marlow, S. L., Lacerenza, C. N., Paoletti, J., Burke, C. S., & Salas, E. (2018). Does team communication represent a one-size-fits-all approach?: A meta-analysis of team communication and performance. *Organizational Behavior and Human Decision Processes*, *144*, 145-170.

Mathieu, J. E. (1991). A cross-level nonrecursive model of the antecedents of organizational commitment and satisfaction. *Journal of Applied Psychology*, *76*(5), 607-618.

Mathieu, J. E., Gallagher, P. T., Domingo, M. A., & Klock, E. A. (2019). Embracing Complexity: Reviewing the Past Decade of Team Effectiveness Research. *Annual Review of Organizational Psychology and Organizational Behaviour*, *6,* 17-46.

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology*, *102*(3), 452-467.

Mathieu, J. E., Kukenberger, M. R., D'innocenzo, L., & Reilly, G. (2015). Modeling reciprocal team cohesion–performance relationships, as impacted by shared leadership and members' competence. *Journal of Applied Psychology*, *100*(3), 713-734.

Mathieu, J. E., Luciano, M. M., & DeChurch, L. A. (2018). Multi-team Systems: The Next Chapter. In N. Anderson, D. Ones, C. Viswesvaran & H. K. Sinangil, (Eds.), *The SAGE Handbook of Industrial, Work & Organizational Psychology*. (pp. 333-353). Sage.

Mathieu, J. E., Marks, M. A., & Zaccaro, S. J. (2001). Multi-team systems. In N. Anderson, D. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *International Handbook of Work and Organizational Psychology*. (pp. 289-313). Sage.

Mathieu, J. E., Wolfson, M. A., & Park, S. (2018). The evolution of work team research since Hawthorne. *American Psychologist*, *73*(4), 308-321.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410-476.

Matusik, J. G., Heidl, R., Hollenbeck, J. R., Yu, A., Lee, H. W., & Howe, M. (2018). Wearable bluetooth sensors for capturing relational variables and temporal variability in relationships: A construct validation study. *Journal of Applied Psychology*, *104*(3), 357-387.

Maynard, M. T., & Kennedy, D. M. (2016). *Team adaptation and resilience: What do we know and what can be applied to long-duration isolated, confined, and extreme contexts*. National Aeronautics and Space Administration.

Maynard, M. T., Kennedy, D. M., & Resick, C. J. (2018). Teamwork in extreme environments: Lessons, challenges, and opportunities. *Journal of Organizational Behaviour*, *39*(6), 695-700.

Maynard, M. T., Kennedy, D. M., & Sommer, S. A. (2015). Team adaptation: A fifteen-year synthesis (1998–2013) and framework for how this literature needs to "adapt" going forward. *European Journal of Work and Organizational Psychology*, *24*(5), 652-677.

Maynard, M. T., Mathieu, J. E., Gilson, L. L., R. Sanchez, D., & Dean, M. D. (2019). Do I really know you and does it matter? Unpacking the relationship between familiarity and information elaboration in global virtual teams. *Group & Organization Management*, *44*(1), 3-37.

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., Lasky, R. E., Sexton, B, J. & Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. *American Journal of Surgery*, *197*(5), 678-685.

McCabe, B., Loughlin, C., Munteanu, R., Tucker, S., & Lam, A. (2008). Individual safety and health outcomes in the construction industry. *Canadian Journal of Civil Engineering*, *35*(12), 1455-1467.

McClurg, C. E., Chen, J. L., Petruzzelli, A., & Thayer, A. L. (2017). Challenges and new directions in examining team cohesion over time. In E. Salas, W. B. Vessey, & L. B Landon (Eds.), *Team Dynamics Over Time.* (pp 261-286). Emerald Publishing Limited.

McEwan, D., & Beauchamp, M. R. (2014). Teamwork in sport: a theoretical and integrative review. *International Review of Sport and Exercise Psychology*, *7*(1), 229-250.

McGrath, J. E. (1964). *Social Psychology: A Brief Introduction*. Holt, Rinehart and Winston.

McGrath, J. E., & Berdahl, J. L. (2002). Groups, technology, and time. In Tindale, R.S., Heath, L., Edwards, J., Posavac, E.J., Bryant, F.B., Myers, J., Suarez-Balcazar, Y., Henderson-King, E. (Eds.) *Theory and Research on Small Groups* (pp. 205-228). Springer.

McMaster, R., & Baber, C. (2012). Multi-agency operations: cooperation during flooding. *Applied Ergonomics*, *43*(1), 38-47.

McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modelling: A Multidisciplinary Journal*, *23*(5), 750-773.

Meehan, M., Insko, B., Whitton, M., & Brooks Jr, F. P. (2002). Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics, 21*(3), 645-652.

Meglino, B. M., & Ravlin, E. C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management*, *24*(3), 351-389.

Mikalachki, A. (1969). Group cohesion revisited. *London, ON: School of Business Administration.*

Militello, L. G., Sushereba, C. E., Branlat, M., Bean, R., & Finomore, V. (2015). Designing for military pararescue: Naturalistic decision-making perspective, methods, and frameworks. *Journal of Occupational and Organizational Psychology, 88*(2), 251-272.

Millward, L. J., Banks, A., & Riga, K. (2010). Effective self-regulating teams: a generative psychological approach. *Team Performance Management: An International Journal*, *16*(1/2), 50-73.

Mishra, J. L., Allen, D. K., & Pearman, A. D. (2011). Information sharing during multi-agency major incidents. *Proceedings of the American Society for Information Science and Technology*, *48*(1), 1-10.

Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, *22*(2), 89-106.

Mohammed, S., Hamilton, K., & Lim, A. (2009). The incorporation of time in team research: Past, current, and future. In E. Salas, G. F. Goodwin & C. S . Burke (Eds.), *Team Effectiveness in Complex Organizations: Cross-disciplinary perspectives and approaches*: (pp. 321–348) Routledge Taylor and Francis.

Molleman, E. (2005). Diversity in demographic characteristics, abilities and personality traits: Do faultlines affect team functioning?. *Group Decision and Negotiation*, *14*(3), 173-193.

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behaviour*, *14*(2), 224-247.

Mozos, O. M., Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N., Dobrescu, R., & Ferrandez, J. M. (2017). Stress detection using wearable physiological and sociometric sensors. *International Journal of Neural Systems*, *27*(2), 1-17.

Murase, T., Carter, D. R., DeChurch, L. A., & Marks, M. A. (2014). Mind the gap: The role of leadership in multiteam system collective cognition. *The Leadership Quarterly*, *25*(5), 972-986.

National Transportation Safety Board (1994). *A Review of Flightcrew-Involved Major Accidents of U.S. Air Carriers, 1978 through 1990*. National Transportation Safety Board.

Nicolas, M., Suedfeld, P., Weiss, K., & Gaudino, M. (2016). Affective, social, and cognitive outcomes during a 1-year wintering in Concordia. *Environment and Behavior*, *48*(8), 1073-1091.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1-13.

O'Neill, T. A., & McLarnon, M. J. (2018). Optimizing team conflict dynamics for high performance teamwork. *Human Resource Management Review*, *28*(4), 378-394.

Okhuysen, G. A. (2001). Structuring change: Familiarity and formal interventions in problem-solving groups. *Academy of Management Journal*, *44*(4), 794-808.

Orasanu, J. & Salas, E. (1993). Team decision making in complex environments. In G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsambok (Eds.), *Decision Making in Action: Models and Methods.* (pp. 327-345). Ablex.

Ortner, S. (1999). *Life and Death on Mount Everest: Sherpas and Himalayan Mountaineering*. Princeton University Press.

Palinkas, L. A. (2003). The psychology of isolated and confined environments: Understanding human behavior in Antarctica. *American Psychologist*, *58*(5), 353-363.

Palinkas, L. A., & Suedfeld, P. (2008). Psychological effects of polar expeditions. *The Lancet*, *371*(9607), 153-163.

Palinkas, L. A., Gunderson, E., Holland, A. W., Miller, C., & Johnson, J. C. (2000). Predictors of behavior and performance in extreme environments: the Antarctic space analogue program. *Aviation, Space, and Environmental Medicine*, *71*(6), 619-625.

Pan, X., & Hamilton, A. F. D. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395-417.

Peeters, M. A., Van Tuijl, H. F., Rutte, C. G., & Reymen, I. M. (2006). Personality and team performance: a meta-analysis. *European Journal of Personality, 20*(5), 377-396.

Pentland, A. (2012). The new science of building great teams. *Harvard Business Review*, *90*(4), 60-69.

Perry, C. W. H. S. J., & Wears, R. L. (2011). Large-scale coordination of work: Coping with complex chaos within healthcare. In K. Mosier & U. M. Fishcer, (Eds,). *Informed by Knowledge* (pp. 69-82). Psychology Press.

Peterson, E., Mitchell, T. R., Thompson, L., & Burr, R. (2000). Collective efficacy and aspects of shared mental models as predictors of performance over time in work groups. *Group Processes & Intergroup Relations*, *3*(3), 296-316.

Pfadenhauer, M. (2009). At eye level: the expert interview—a talk between expert and quasi-expert. In A. Bogner, B. Littig & W. Menz (Eds.), *Interviewing Experts* (pp. 81-97). Palgrave Macmillan.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903.

Pollock, K. (2017) *Local Interoperability in UK Emergency Management: A Research Report.* Emergency Planning College.

Porter, C. O. L. H., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P. J., West, B. J., & Moon, H. (2003). Backing up behaviors in teams: The role of personality and legitimacy of need. *Journal of Applied Psychology, 88*(3), 391–403.

Power, N. (2015). *Cognition in Crisis: Decision Inertia and Failures to take Action in Multi-Agency Emergency Response Command Teams* (Unpublished doctoral thesis). University of Liverpool, Liverpool.

Power, N. (2018). Extreme teams: Toward a greater understanding of multiagency teamwork during major emergencies and disasters. *American Psychologist*, *73*(4), 478-490.

Power, N., & Alison, L. (2017a). Redundant deliberation about negative consequences: Decision inertia in emergency responders. *Psychology, Public Policy, and Law*, *23*(2), 243-258.

Power, N., & Alison, L. (2017b). Offence or defence? Approach and avoid goals in the multi-agency emergency response to a simulated terrorism attack. *Journal of Occupational and Organizational Psychology*, *90*(1), 51-76.

Qi, Y., Zhao, X., & Sheu, C. (2011). The impact of competitive strategy and supply chain strategy on business performance: the role of environmental uncertainty. *Decision Sciences*, *42*(2), 371-389.

Quarantelli, E. L. (1985). Organisational behaviour in disasters and implications for disaster planning. *Monographs of the National Emergency Training Center, 1*(2), 1-31.

Reagans, R., Argote, L., & Brooks, D. (2005). Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science*, *51*(6), 869-881.

Ren, Y., & Argote, L. (2011). Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals*, *5*(1), 189-229.

Rico, R., Hinsz, V. B., Burke, S., & Salas, E. (2017). A multilevel model of multiteam motivation and performance. *Organizational Psychology Review*, *7*(3), 197-226.

Rico, R., Hinsz, V. B., Davison, R. B., & Salas, E. (2018). Structural influences upon coordination and performance in multiteam systems. *Human Resource Management Review*, *28*(4), 332-346.

Risser, D. T., Rice, M. M., Salisbury, M. L., Simon, R., Jay, G. D., Berns, S. D., & MedTeams Research Consortium. (1999). The potential for improved teamwork to reduce medical errors in the emergency department. *Annals of Emergency Medicine*, *34*(3), 373-383.

Roberto, M. A. (2002). Lessons from Everest: The interaction of cognitive bias, psychological safety, and system complexity. *California Management Review*, *45*(1), 136-158.

Roberts, K. H. (1990). Managing high reliability organizations. *California Management Review*, *32*(4), 101-113.

Roma, P. G., & Bedwell, W. L. (2017). Key factors and threats to team dynamics in long-duration extreme environments. In E. Salas, W. B. Vessey, & L. B Landon (Eds.), *Team Dynamics Over Time* (pp. 155-187). Emerald Publishing Limited.

Rosen, M. A., & Dietz, A. S. (2017). Team performance measurement. In E. Salas, R. Rico & J. Passmore, (Eds.), *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*, (pp. 479-502). Wiley Blackwell.

Rosen, M. A., Fiore, S. M., Salas, E., Letsky, M., & Warner, N. (2008). Tightly coupling cognition: understanding how communication and awareness drive coordination in teams. *International Journal of Command and Control, 2*(1), 1–30.

Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S., Robinson, D. & Birnbach, D. J. (2008). Measuring team performance in simulation-based training: adopting best practices for healthcare. *Simulation in Healthcare*, *3*(1), 33-41.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modelling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1-36.

Rousseau, D. M., & Fried Y. (2001). Location, location, location: Contextualizing organizational research. *Journal of Organizational Behaviour, 22*(1), 1–13.

Rousseau, V., Aubé, C., & Savoie, A. (2006). Teamwork behaviors: A review and an integration of frameworks. *Small Group Research*, *37*(5), 540-570.

Rudolph, J. W., Simon, R., & Raemer, D. B. (2007). Which reality matters? Questions on the path to high engagement in healthcare simulation. *Simulation in Healthcare, 2*(3), 161-163.

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*(1), 471-499.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors*, *50*(3), 540-547.

Salas, E., DiazGranados, D., Weaver, S. J., & King, H. (2008). Does team training work? Principles for health care. *Academic Emergency Medicine*, *15*(11), 1002-1009.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance*. Ablex.

Salas, E., Driskell, J.E., & Hughes, S. (1996). Introduction: The study of stress and human performance. In J.E. Driskell & E. Salas (Eds.), *Stress and Human Performance* (pp. 1-46). Erlbaum.

Salas, E., Goodwin, G. F., & Burke, C. S. (2008). *Team Effectiveness in Complex Organizations: Cross-disciplinary Perspectives and Approaches*. Routledge.

Salas, E., Grossman, R., Hughes, A. M., & Coultas, C. W. (2015). Measuring team cohesion: Observations from the science. *Human Factors*, *57*(3), 365-374.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors, 37*(1), 123–136.

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist*, *73*(4), 593-600.

Salas, E., Rico, R., & Passmore, J. (2017). *The Wiley Blackwell Handbook of the Psychology of Team working and Collaborative Processes*. Wiley Blackwell.

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management*, *54*(4), 599-622.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork?. *Small Group Research*, *36*(5), 555-599.

Salas, E., Sims, D. E., & Klein, C. (2004). Cooperation at Work. In C. D. Speilberger (Ed.), *Encyclopedia of Applied Psychology.* (pp. 497-507). Academic Press.

Salas, E., Tannenbaum, S. I., Kozlowski, S. W., Miller, C. A., Mathieu, J. E., & Vessey, W. B. (2015). Teams in space exploration: A new frontier for the science of team effectiveness. *Current Directions in Psychological Science*, *24*(3), 200-207.

Salmon, P., Stanton, N., Jenkins, D., & Walker, G. (2011). Coordination during multi-agency emergency response: issues and solutions. *Disaster Prevention and Management: An International Journal*, *20*(2), 140-158.

Sambucini, V. (2017). Bayesian vs Frequentist Power Functions to Determine the Optimal Sample Size: Testing One Sample Binomial Proportion Using Exact Methods. In J.P. Tejedor (Eds.) *Bayesian Inference* (pp 77-97). InTech.

Sandal G, Vaernes R, Ursin H. (1995) Interpersonal relations during simulated space missions. *Aviation, Space, Environmental Medicine*, *66*(7), 617–24.

Sandal, G. M., Bye, H. H., & van de Vijver, F. J. (2011). Personal values and crew compatibility: Results from a 105-days simulated space mission. *Acta Astronautica*, *69*(3-4), 141-149.

Sandal, G. M., Endresen, I. M., Vaernes, R., & Ursin, H. (1999). Personality and coping strategies during submarine missions. *Military Psychology*, *11*(4), 381-404.

Sandal, G. M., Leon, G. R., & Palinkas, L. (2006). Human challenges in polar and space environments. *Reviews in Environmental Science and Bio/Technology, 5,* 281-296.

Sandal, G. M., Smith, N., & Leon, G. R. (2017). Analog Mission Research. In T. Sgobba (Eds.)*, Space Safety and Human Performance,* (pp. 262). Elsevier.

Schmutz, J. B., Lei, Z., Eppich, W. J., & Manser, T. (2018). Reflection in the heat of the moment: The role of in‐action team reflexivity in health care emergency teams. *Journal of Organizational Behaviour*, *39*(6), 749-765.

Schmutz, J., & Manser, T. (2013). Do team processes really have an effect on clinical performance? A systematic literature review. *British Journal of Anaesthesia*, *110*(4), 529-544.

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379-386.

Schraagen, J. M., & van de Ven, J. (2011). Human factors aspects of ICT for crisis management. *Cognition, Technology & Work*, *13*(3), 175-187.

Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology*, *5*(2-3), 137-182.

Schwartz, S. H. (2011). Values: Cultural and individual. In S. M. Breugelmans, A. Chasiotis, & F. J. R. van de Vijver (Eds.), *Fundamental Questions in Cross-cultural Psychology* (pp. 463–493). Cambridge University Press.

Scott, J. (2000). *Social network analysis.* Sage.

Seashore, S. E. (1954). *Group Cohesiveness in the Industrial Work Group*. Institute for Social Research.

Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communications: the links between language, performance, error, and workload. *Human Performance in Extreme Environments*, *5*(1), 63-68.

Shah, P. P., & Jehn, K. A. (1993). Do friends perform better than acquaintances? The interaction of friendship, conflict, and task. *Group Decision and Negotiation*, *2*(2), 149-165.

Shuffler, M. L., & Carter, D. R. (2018). Teamwork situated in multiteam systems: Key lessons learned and future opportunities. *American Psychologist*, *73*(4), 390-406.

Shuffler, M. L., Jiménez-Rodríguez, M., & Kramer, W. S. (2015). The science of multiteam systems: A review and future research agenda. *Small Group Research*, *46*(6), 659-699.

Sieweke, J., & Zhao, B. (2015). The impact of team familiarity and team leader experience on team coordination errors: A panel analysis of professional basketball teams. *Journal of Organizational Behaviour*, *36*(3), 382-402.

Sleeper, J. A., & Thompson, C. (2008). The use of hi fidelity simulation to enhance nursing students' therapeutic communication skills. *International Journal of Nursing Education Scholarship*, *5*(1), 1-12.

Smith-Jentsch, K. A., Kraiger, K., Cannon-Bowers, J. A., & Salas, E. (2009). Do familiar teammates request and accept more backup? Transactive memory in air traffic control. *Human Factors*, *51*(2), 181-192.

Smith, N., & Barrett, E. C. (2019). Psychology, extreme environments, and counter-terrorism operations. *Behavioural Sciences of Terrorism and Political Aggression*, *11*(1), 48-72.

Smith, N., Barrett, E., & Sandal, G. M. (2018). Monitoring daily events, coping strategies, and emotion during a desert expedition in the Middle East. *Stress and Health*, *34*(4), 534-544.

Smith, N., Kinnafick, F., & Saunders, B. (2017). Coping strategies used during an extreme Antarctic expedition. *Journal of Human Performance in Extreme Environments*, *13*(1), 1.

Smith, N., Kinnafick, F., Cooley, S. J., & Sandal, G. M. (2017). Reported growth following mountaineering expeditions: The role of personality and perceived stress. *Environment and Behaviour*, *49*(8), 933–955.

Smith, W., & Dowell, J. (2000). A case study of co-ordinative decision-making in disaster management. *Ergonomics*, *43*(8), 1153-1166.

Sneddon, A., Mearns, K., & Flin, R. (2006). Situation awareness and safety in offshore drill crews. *Cognition, Technology & Work, 8*(4), 255-267

Soda, G., Usai, A., & Zaheer, A. (2004). Network memory: The influence of past and current networks on performance. *Academy of Management Journal*, *47*(6), 893-906.

Staats, B. R., Gino, F., & Pisano, G. P. (2010). Varied experience, team familiarity, and learning: The mediating role of psychological safety. Working paper 10-016, Harvard Business School.

Stachowski, A. A., Kaplan, S. A., & Waller, M. J. (2009). The benefits of flexible team interaction during crises. *Journal of Applied Psychology*, *94*(6), 1536-1543.

Stanton, N. A., & Roberts, A. P. J. (2018). Examining social, information, and task networks in submarine command and control. *IEEE Transactions on Human-Machine Systems*, *48*(3), 252-265.

Steel, G. D., Suedfeld, P., Peri, A., & Palinkas, L. A. (1997). People in high latitudes: the" Big Five" personality characteristics of the circumpolar sojourner. *Environment and Behavior*, *29*(3), 324-347.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors*, *41*(1), 61-71.

Stuster, J. (2011). *Bold Endeavours: Lessons from Polar and Space Exploration*. Naval Institute Press.

Suedfeld, P., & Steel, G. D. (2000). The environmental psychology of capsule habitats. *Annual Review of Psychology*, *51*(1), 227-253.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1-39.

Taylor, P. J., (June, 2013). How technology is revolutionizing our understanding of human cooperation (Inaugural lecture). Twente University Press. Downloaded on 20 June, 2019, from https://www.utwente.nl/en/academic-ceremonies/inaugural-lectures/booklets-inaugural-lectures/2007-2014/Oratieboekje_Taylor.pdf

Turner, J. R., Thurlow, N., Baker, R., Northcutt, D., & Newman, K. (2019). Multiteam systems in an agile environment: a realist systematic review. *Journal of Manufacturing Technology Management*, *30*(4), 748-771.

Uitdewilligen, S., & Waller, M. J. (2018). Information sharing and decision-making in multidisciplinary crisis management teams. *Journal of Organizational Behaviour*, *39*(6), 731-748.

Valentine, M. A., Nembhard, I. M., & Edmondson, A. C. (2015). Measuring teamwork in health care settings: a review of survey instruments. *Medical Care*, *53*(4), 16-30.

Van der Haar, S., Segers, M., Jehn, K., & Van den Bossche, P. (2015). Investigating the relation between team learning and the team situation model. *Small Group Research*, *46*(1), 50-82.

Van Vianen, A. E., & De Dreu, C. K. (2001). Personality in teams: Its relationship to social cohesion, task cohesion, and team performance. *European Journal of Work and Organizational Psychology*, *10*(2), 97-120.

Vandekerckhove, J., Rouder, J. N., & Kruschke, J. K. (2018). Bayesian methods for advancing psychological science. *Psychonomic Bulletin and Review, 25*(1), 1-4.

Vessey, W. B., & Landon, L. B. (2017). Team Performance in Extreme Environments. In In E. Salas, R. Rico & J. Passmore, (Eds.), *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*, (pp. 531-553). Wiley Blackwell.

Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. *The Journal of Applied Behavioural Science*, *41*(4), 373-398.

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q, F., Smira, M., Epskamp, S., Matzke, D., Rouder, J. N., Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35-57.

Wagstaff, C. R., & Weston, N. J. (2014). Examining emotion regulation in an isolated performance team in Antarctica. *Sport, Exercise, and Performance Psychology*, *3*(4), 273-287.

Walker, G. H., Gibson, H., Stanton, N. A., Baber, C., Salmon, P., & Green, D. (2006). Event analysis of systemic teamwork (EAST): a novel integration of ergonomics methods to analyse C4i activity. *Ergonomics*, *49*(12-13), 1345-1369.

Walker, G. H., Stanton, N. A., Baber, C., Wells, L., Gibson, H., Salmon, P., & Jenkins, D. (2010). From ethnography to the EAST method: A tractable approach for representing distributed cognition in Air Traffic Control. *Ergonomics*, *53*(2), 184-197.

Wallenius, C., Larsson, G., & Johansson, C. R. (2004). Military observers' reactions and performance when facing danger. *Military Psychology*, *16*(4), 211-229.

Waller, M. J., & Kaplan, S. A. (2018). Systematic behavioural observation for emergent team phenomena: Key considerations for quantitative video-based approaches. *Organizational Research Methods*, *21*(2), 500-515.

Waller, M. J., Gupta, N., & Giambatista, R. C. (2004). Effects of adaptive behaviors and shared mental models on control crew performance. *Management Science*, *50*(11), 1534-1544.

Waller, M.J., Utidewilligen S. (2008) Talking to the room: collective sensemaking during crisis situations. In R. Roe, M. J. Waller & S. Clegg (Eds.), *Time in Organizations—Approaches and Methods.* (pp. 186-204). Routledge.

Wang, M., & Hanges, P. J. (2011). Latent class procedures: Applications to organizational research. *Organizational Research Methods*, *14*(1), 24-31.

Waring, S., Alison, L., Carter, G., Barrett-Pink, C., Humann, M., Swan, L., & Zilinsky, T. (2018). Information sharing in interteam responses to disaster. *Journal of Occupational and Organizational Psychology*, *91*(3), 591-619.

Waring, S., Alison, L., Shortland, N., & Humann, M. (2019). The role of information sharing on decision delay during multiteam disaster response. *Cognition, Technology & Work*, https://doi.org/10.1007/s10111-019-00570-7

Waring, S., Moran, J., Page, R. (2020). Decision-making in multiagency multiteam systems operating in extreme environments. *Journal of Occupational and Organizational Psychology.* https://doi.org/10.1111/joop.12309.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and applications*. Cambridge University Press.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063-1070.

Wauben, L. S. G. L., Dekker-van Doorn, C. M., Van Wijngaarden, J. D. H., Goossens, R. H. M., Huijsman, R., Klein, J., & Lange, J. F. (2011). Discrepant perceptions of communication, teamwork and situation awareness among surgical team members. *International Journal for Quality in Health Care*, *23*(2), 159-166.

Waugh, William L., Jr. (2004). Leveraging Networks to Meet National Goals: FEMA and the Safe Construction Networks. In J. M. Kamensky & T. J. Burlin (Eds.) *Collaboration: Using Partnerships and Networks,* (pp. 273-320). Rowman & Littlefield Publishers.

Weaver, S. J., Dy, S. M., & Rosen, M. A. (2014). Team-training in healthcare: a narrative synthesis of the literature. *BMJ Quality* Safety, *23*(5), 359-372.

Wei, W., Pfeffer, J., Reminga, J., & Carley, K. M. (2011). Handling weighted, asymmetric, self-looped, and disconnected networks in ORA. *SSRN.* http://dx.doi.org/10.2139/ssrn.2729311

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, *38*(3), 357-381.

Weiss, K., & Gaud, R. (2004). Formation and Transformation of Relational Networks During an Antarctic Winter-Over. *Journal of Applied Social Psychology*, *34*(8), 1563-1586.

Weller, J. M., Janssen, A. L., Merry, A. F., & Robinson, B. (2008). Interdisciplinary team interactions: a qualitative study of perceptions of team function in simulated anaesthesia crises. *Medical Education*, *42*(4), 382-388.

West, C., Landry, K., Graham, A., Graham, L., Cianciolo, A. T., Kalet, A., Rosen, M. & Sherman, D. W. (2015). Conceptualizing interprofessional teams as multi-team systems—Implications for assessment and training. *Teaching and Learning in Medicine*, *27*(4), 366-369.

Westli, H. K., Johnsen, B. H., Eid, J., Rasten, I., & Brattebø, G. (2010). Teamwork skills, shared mental models, and performance in simulated trauma teams: an independent group design. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *18*(1), 47.

Wetzel, C. M., Kneebone, R. L., Woloshynowych, M., Nestel, D., Moorthy, K., Kidd, J., & Darzi, A. (2006). The effects of stress on surgical performance. *The American Journal of Surgery*, *191*(1), 5-10.

Wickens, C. D., Keller, J. W., & Shaw, C. (2015). Human factors in high-altitude mountaineering. *Journal of Human Performance in Extreme Environments*, *12*(1), 1.

Wijnmaalen, J., Voordijk, H., Rietjens, B. (2018). MTS coordination in practice: micro-level insights to increase MTS performance. *Team Performance Management, 24(*1/2), 64–83.

Wijnmaalen, J., Voordijk, H., Rietjens, S., & Dewulf, G. (2019). Intergroup behavior in military multiteam systems. *Human Relations*, *72*(6), 1081-1104.

Wilkinson, B., Cohen-Hatton, S. R., & Honey, R. C. (2019). Decision-making in multi-agency groups at simulated major incident emergencies: In situ analysis of adherence to UK doctrine. *Journal of Contingencies and Crisis Management*, *27*(1), 306-316.

Willemsen-Dunlap, A. M., Binstadt, E. S., Nguyen, M. C., Elliott, N. C., Cheney, A. R., Stevens, R. H., & Dooley-Hash, S. (2018). Alternative Markers of Performance in

Simulation: Where We Are and Where We Need To Go. *Academic Emergency Medicine*, *25*(2), 250-254.

Wood, J., Schmidt, L., Lugg, D., Ayton, J., Phillips, T., & Shepanek, M. (2005). Life, survival, and behavioral health in small closed communities: 10 years of studying isolated Antarctic groups. *Aviation, Space, and Environmental Medicine*, *76*(6), 89-93.

Wu, L., Waber, B., Aral, S., Brynjolfsson, E., & Pentland, A. (2008). Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an it configuration task. Retrieved from: SSRN. http://dx.doi.org/10.2139/ssrn.1130251

Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes*, *118*(2), 179-188.

Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., & Paterson-Brown, S. (2008). Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World Journal of Surgery*, *32*(4), 548-556.

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., & Rowley, D. (2006). Development of a rating system for surgeons' non-technical skills. *Medical Education*, *40*(11), 1098-1104.

Zaccaro, S. J., Gualtieri, J., & Minionis, D. (1995). Task cohesion as a facilitator of team decision making under temporal urgency. *Military Psychology*, *7*(2), 77-93.

Zaccaro, S., Marks, M., & DeChurch, L. (2012). *Multiteam Systems: An Organizational Form for Dynamic and Complex Environments.* Routledge Academic.

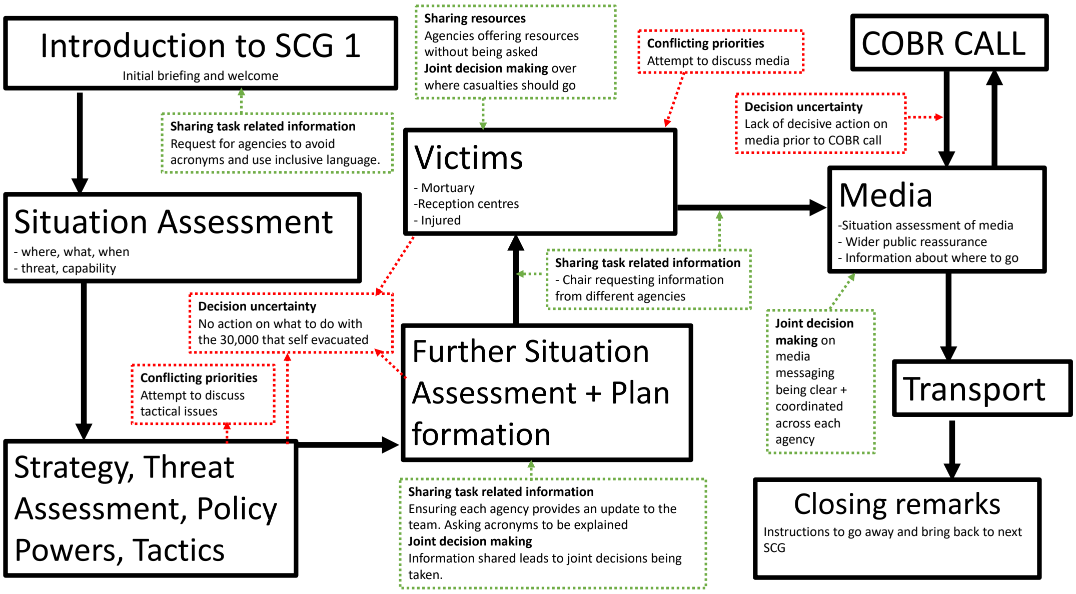
Zhang, Y., Olenick, J., Chang, C. H., Kozlowski, S. W., & Hung, H. (2018). TeamSense: Assessing Personal Affect and Group Cohesion in Small Teams through Dyadic Interaction and Behavior Analysis with Wearable Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 150:1-22.

Zhou, Y., & Wang, E. (2010). Shared mental models as moderators of team process-performance relationships. *Social Behaviour and Personality, 38*(4), 433-444.

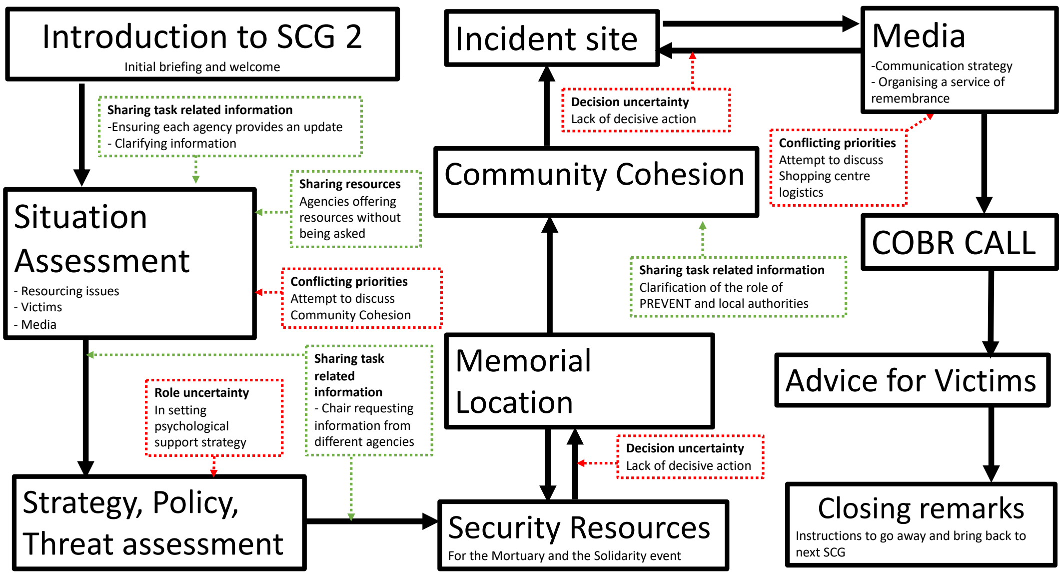
Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, *41*(2), 390-420.

Appendices

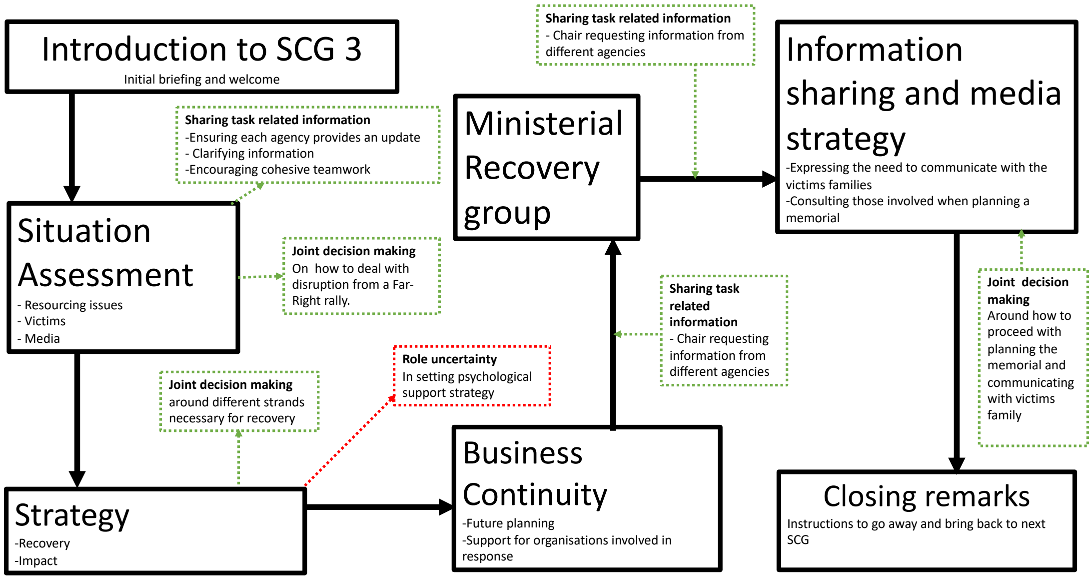
Appendix 1 Task flow for each SCG



Appendix 1.1 Task flow of SCG1



Appendix 1.2 Task flow of SCG2



Appendix 1.3 Task flow of SCG3

Appendix 2

Coding framework for coordinating behaviours

Top-level codes retained for analysis	**Codes**	**Examples**
Joint decision making	Requesting input from others for decision making	"Okay, I take the assessment from you, as to whether trying to close the M1 and the gridlock around that and displacing traffic or whether we leave it free flowing and just getting on there with convoy" (pg. 12 SCG1)
	Encouraging joint decision making	"If we could just look strategically then and just to be jointly clear, whatever you do, don't rely on me seeing all the areas of vulnerability. I'm not going to spot half of them. This very much is a team effort" (pg. 15 SCG1) "I think the more that we can be seen as cohesively as a team, acting and working together, appearing together, the better that presents itself for all concerned." (pg. 8, SCG2)
	Joint decision making	"I've spoken with the coroner in terms of suspect identification, to influence that wider CT investigative network." (pg. 5, SCG2) "Are we agreed in that case on that as a strategy" (pg. 11, SCG3)
Sharing Resources	Offering assistance	"In terms of government support, I'm happy to facilitate anything in terms of speeding up the military assistance and contacting ministers." (pg. 9 SCG1) "Chair can I just add, that I've just had confirmation that our volunteers have all been stood up and ready to assist in whatever way that we are asked to." (pg. 13 SCG1) "Okay, so would you be able to, on the basis of your experience in the past be able to draw up some sort of strategy for us in terms of a hierarchy of how we distribute that." (pg. 7, SCG 2)

		“Would it be helpful to have red cross volunteers there, possibly street pastors, crews, on hand to talk to people if they find it emotionally challenging?” (pg. 14, SCG3)
Sharing task related information	Requesting information by chair	“Okay, before we start then is there any urgent business to attend to right at the start?” (pg. 2 SCG2). “Could I possibly ask XX Police to give an update as to where we are to begin with?” (pg. 2 SCG3)
	Overview provided by chair	“Quickly moving through the strategy, just to bring a snapchat of rationale behind each of those so we are all clear on the basis, that we move forward from” (pg. 2 SCG1)
	Explaining acronyms	“If a health colleague could just explain the p2,p3 and p1, just so we are all aware of that terminology please.” (pg. 3 SCG1).
Role uncertainties	A lack of role understanding	“So, I welcome the offer from the Red cross but its important that we take that through the mental health networks in XX, so that we have a coordinated response.” (pg. 11, SCG2) “Can I just check in terms of that.. I think the health responsibility remains with the Health, so just to be clear on that.. we will draw on colleages from local authorities” (pg. 11, SCG3)
Decision uncertainties	Delay in response due to missed information or indecision	“There is a high demand for information, that we are not really able to respond to at this moment in time.” (pg. 11, SCG1) “Can I express some concern over the comms vacuum at the moment that we do have, in terms of longevity of time.” (pg. 27, SCG1) “Okay, so deal with this issue of strategic police reserve. I’m just conscious that we need to deal with this issue of memorial team. So its Barkers.. Barkers Pool.”(pg. 15, SCG2)
Conflicting priorities	Distracting from the task at hand	“So if we deal with the victims first, where they are going to get that out and then part of that once we’ve got clarity will be our internal external comms approach, so very shortly we will get to comms.” (pg. 16 SCG1)

		"I think what we will do is we will deal with community cohesion as a separate agenda item. If we just do an update on what your current issues, what the current SITREP is, for all of you." (pg. 7 SCG2)
		"just in terms of.. trying to keep it as at the strategic level as I can" (pg. 12 SCG2)
	Priority raised and not dealt with	"Well I think we need to.. I think we need to work together to actually plan into that later." (pg. 7 SCG2)
		"Okay, I think what id do with that, is again task that on, not deal with that here in the SCG," (pg. 19, SCG2)

Appendix 3

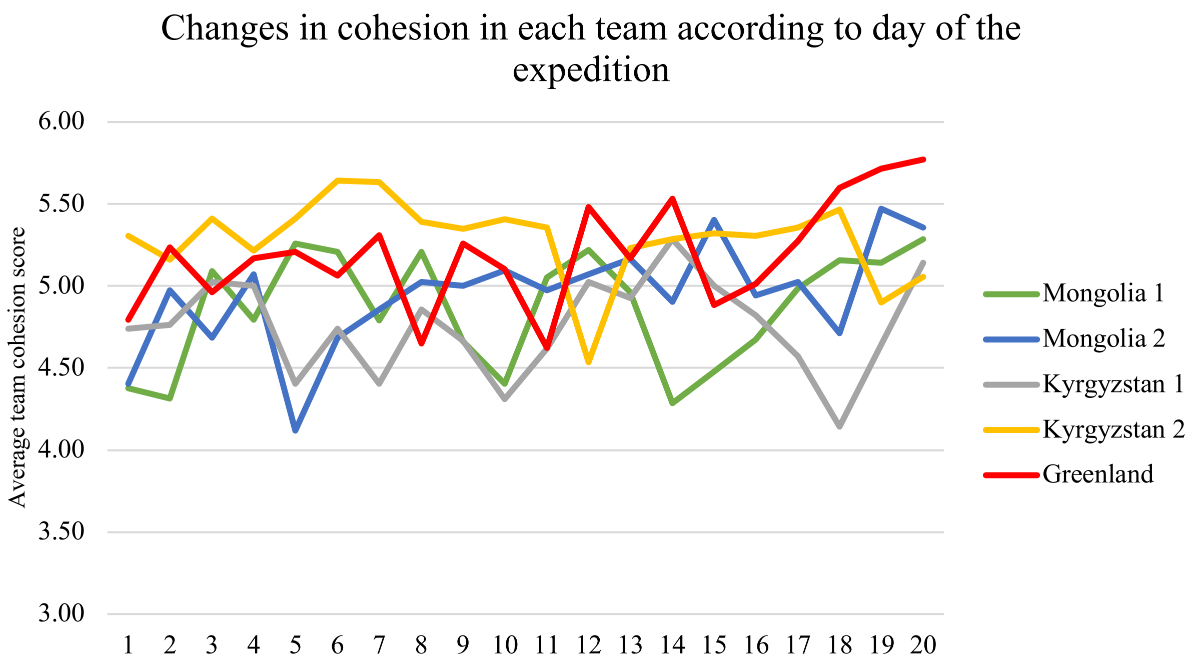
Thematic analysis of coordinating behaviours using Brown & Power (2018) codebook.

Coordinating Behaviour	Description	Examples
Joint decision making	Any instance in which the team actively worked together to implement a decision	"Does everyone agree that this is too close?" (Team 1). "I'm just going to tally if anyone said anything similar to mine and then we can just have an idea at the end" (Team 11) "So, I think we've gained consensus on three main priorities there." (Team 13) "A? Anyone? I think A, but we've got three Cs. B? Anyone?" (Team 14)
Sharing resources	Any instance in which team members offered resources/support to assist other agencies.	"I mean, if it comes to fire, I don't think you need more than 16, but the Fire and Rescue needs to help you in getting people out, it is not just about getting the fire down" (Team 3) "We have voluntary agencies around the area, so we can send those to the reception area." (Team 13) "We'll have officers there to make sure the press don't enter the university campus" (Team 16) "And I have the Red Cross who want to go in and help" (Team 22)
Sharing task related information	Any instance in which team members actively sought to improve communication and a shared situational awareness.	"I have four-armed response vehicles, each having six personnel, so I can send them in." (Team 1) "Yeah. So, there's 100 plus casualties and I can fit 60 in my nearby hospitals" (Team 7) "I'd say about 20,000 runners signed up to the marathon and the marathon ends here, so then they go through there" (Team 18)

		"I would say send in the SAR, which is essentially like a team who are like specifically designed to work in hazardous environments" (Team 19)
Role uncertainties	Confusion about one's own role and the role of others.	"Only the search and rescue are the ones that are trained and have protective gear so those are the only people that are allowed in the hot zone." (Team 4) "So, as a mayor, also my responsibility would be to direct all the agencies" (Team 6) "Yeah. I don't really think this is too relevant to me, I think I just need to give you guys information about what's happening" (Team 16) "Take responders out, get Fire and Rescue to sweep for bombs and everything, keep the police line solid and keep everyone out of the area"(Team 20)
Decision uncertainties	Any instance in which decision-making lack clarity and discussions were delayed due to indecision.	" I don't know what's going on" (Team 7) "So, I'm confused as to what to do" (Team 6) "Wait, did the attack happen during a marathon?" (Team 15) "What information are we waiting around for?" (Team 19)
Conflicting priorities	Any instance in which team members attempted to re-orient the conversation towards intra-agency priorities.	"I think risks is different, it's a different thing, but okay." (Team 13) "That is going to take too long, I think" (Team 3). "It's not a priority, is it?" (Team 11) "Is that really important right now, though?" (Team 22

Appendix 4. Changes in cohesion over the course of the expedition

Appendix 4 was not included in the original published paper due to the page limit, however, it is useful to illustrate the changes in cohesion in each team over time. As identified in the Figure below, cohesion fluctuates differently in each team, suggesting that it does not simply emerge in a linear manner but is influenced by (as demonstrated in Table 2, *Chapter V*), the dynamics of the team and the external environment.



Appendix 5. Revised Diary

Date ………………….

Please provide a brief summary of what you achieved today (if possible please reflect on the goal that you set yourself yesterday)

Please tick each event that you experienced today

Daily event		**Daily event**	
Problems with gear and equipment e.g., clothing, tools, communication, navigation equipment etc.		Feeling I could rely on my team mates to work effectively	
A delay in progress due to weather conditions		My team mates approached the expedition today with a good attitude	
Worried about encountering bad weather		Satisfaction with the leadership	
Enjoyment of the environment		A problem/dispute with the leader	
Satisfaction that equipment is working properly		Problems with digestion	
Satisfaction in making good progress today		Lack of sleep	
Satisfaction that I am able to cope with challenges		Muscle, joint or injury (incl. headache,	
Concerns about the effectiveness or safety of the decisions I made today		Personal hygiene (wanting to be cleaner)	
Concern about the well-being of my team mates		Lack of privacy, personal time	
Tension or argument with my team mate(s)		Fear of being injured	
Feeling down/low stressed out because my team mates is/are feeling that way		Loneliness, homesickness	
Feeling of camaraderie/closeness with team mates		Worried about family/friends	
Concern about how effectively my team mates and I are working together		Feeling of monotony/boredom	

Please indicate your agreement with the following statements about your team

	Completely disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Completely agree
A feeling of unity and cohesion in my team							
A strong feeling of belongingness among my team.							
Members of my team felt close to each other							
My team failed to adopt a shared a focus on our work							
My team concentrated on getting things done							
Team members did not pull together to accomplish the work							
Overall, my team performed effectively today							
I think I performed well in the fulfillment of my individual duties							

Please evaluate these statements regarding relatedness to your family and friends

1 = Strongly agree **2** = moderately agree **3 =** neither agree nor disagree **4 =** moderately disagree **5 =** strongly disagree

1. I felt excluded from my friends and family at home ____
2. I felt connected to my friends and family at home, who likewise cared about me ____
3. I felt my friends and family at home acting distant towards me. ____
4. I had a warm feeling about my friends and family at home. ____

Indicate how you have felt today by putting a number next to each of the emotions

1 = not at all **2** = a little **3 =** moderately **4 =** quite a bit **5 =** extremely

Interested	_____	Guilty	_____	Irritable	_____	Determined	_____
Distressed	_____	Scared	_____	Alert	_____	Attentive	_____
Excited	_____	Hostile	_____	Ashamed	_____	Jittery	_____
Upset	_____	Enthusiastic	_____	Inspired	_____	Active	_____
Strong	_____	Proud	_____	Nervous	_____	Afraid	_____

What is your main goal for tomorrow?